

Discrete Convexity and Polynomial Solvability in Minimum 0-Extension Problems

Hiroshi HIRAI

Department of Mathematical Informatics,
Graduate School of Information Science and Technology,
The University of Tokyo, Tokyo, 113-8656, Japan.
hirai@mist.i.u-tokyo.ac.jp

October, 2012

May, 2014 (revised)

September, 2014 (final)

Abstract

A 0-extension of graph Γ is a metric d on a set V containing the vertex set V_Γ of Γ such that d extends the shortest path metric of Γ and for all $x \in V$ there exists a vertex s in Γ with $d(x, s) = 0$. The minimum 0-extension problem **0-Ext** $[\Gamma]$ on Γ is: given a set $V \supseteq V_\Gamma$ and a nonnegative cost function c defined on the set of all pairs of V , find a 0-extension d of Γ with $\sum_{xy} c(xy)d(x, y)$ minimum. The 0-extension problem generalizes a number of basic combinatorial optimization problems, such as minimum (s, t) -cut problem and multiway cut problem.

Karzanov proved the polynomial solvability of **0-Ext** $[\Gamma]$ for a certain large class of modular graphs Γ , and raised the question: What are the graphs Γ for which **0-Ext** $[\Gamma]$ can be solved in polynomial time? He also proved that **0-Ext** $[\Gamma]$ is NP-hard if Γ is not modular or not orientable (in a certain sense).

In this paper, we prove the converse: if Γ is orientable and modular, then **0-Ext** $[\Gamma]$ can be solved in polynomial time. This completes the classification of graphs Γ for which **0-Ext** $[\Gamma]$ is tractable. To prove our main result, we develop a theory of discrete convex functions on orientable modular graphs, analogous to discrete convex analysis by Murota, and utilize a recent result of Thapper and Žitný on valued CSP.

1 Introduction

By a *(semi)metric* d on a finite set V we mean a nonnegative symmetric function on $V \times V$ satisfying $d(x, x) = 0$ for all $x \in V$ and the triangle inequalities $d(x, y) + d(y, z) \geq d(x, z)$ for all $x, y, z \in V$. An *extension* of a metric space (S, μ) is a metric space (V, d) with $V \supseteq S$ and $d(s, t) = \mu(s, t)$ for $s, t \in S$. An extension (V, d) of (S, μ) is called a *0-extension* if for all $x \in V$ there exists $s \in S$ with $d(s, x) = 0$.

Let Γ be a simple connected undirected graph with vertex set V_Γ . Let d_Γ denote the shortest path metric on V_Γ with respect to the uniform unit edge-length of Γ . The *minimum 0-extension problem* **0-Ext** $[\Gamma]$ on Γ is formulated as:

0-Ext $[\Gamma]$: Given $V \supseteq V_\Gamma$ and $c : \binom{V}{2} \rightarrow \mathbf{Q}_+$,

minimize $\sum_{xy \in \binom{V}{2}} c(xy)d(x, y)$ over all 0-extensions (V, d) of (V_Γ, d_Γ) .

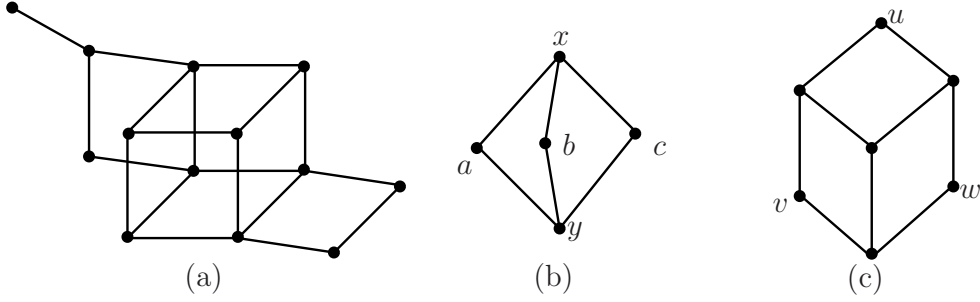


Figure 1: (a) a median graph, (b) a, b, c have two medians x, y , and (c) u, v, w have no median

Here $\binom{V}{2}$ denotes the set of all pairs of V . The minimum 0-extension problem is formulated by Karzanov [32], and is equivalent to the following classical facility location problem, known as *multifacility location problem* [55], where we let $V \setminus V_\Gamma := \{1, 2, \dots, n\}$:

$$(1.1) \quad \begin{aligned} \text{Min.} \quad & \sum_{s \in V_\Gamma} \sum_{1 \leq j \leq n} c(sj) d_\Gamma(s, \rho_j) + \sum_{1 \leq i < j \leq n} c(ij) d_\Gamma(\rho_i, \rho_j) \\ \text{s.t.} \quad & \rho = (\rho_1, \rho_2, \dots, \rho_n) \in V_\Gamma \times V_\Gamma \times \dots \times V_\Gamma. \end{aligned}$$

This problem can be interpreted as follows: We are going to locate n new facilities $1, 2, \dots, n$ on graph Γ , where the facilities communicate each other and communicate existing facilities on Γ . The cost of the communication is proportional to the distance. Our goal is to find a location of minimum total communication cost. This classic facility location problem arises in many practical situations such as the image segmentation in computer vision, and related clustering problems in machine learning; see [36]. Also **0-Ext** $[\Gamma]$ includes a number of basic combinatorial optimization problems. For example, take as Γ the graph K_2 consisting of a single edge st . Then **0-Ext** $[K_2]$ is the minimum (s, t) -cut problem. More generally, **0-Ext** $[K_m]$ is the multiway cut problem on m terminals. Therefore **0-Ext** $[K_m]$ is solvable in polynomial time if $m = 2$ and is NP-hard if $m > 2$ [14].

This paper addresses the following problem considered by Karzanov [32, 34, 35].

*What are the graphs Γ for which **0-Ext** $[\Gamma]$ is solvable in polynomial time?*

Here such a graph is simply called *tractable*.

A classical result in location theory in the 1970's is:

Theorem 1.1 ([51]; also see [37]). *If Γ is a tree, then **0-Ext** $[\Gamma]$ is solvable in polynomial time.*

The tractability of graphs Γ is preserved under taking Cartesian products. Therefore, cubes, grid graphs, and the Cartesian product of trees are tractable. Chepoi [12] extended this classical result to median graphs as follows. A *median* of a triple p_1, p_2, p_3 of vertices is a vertex m satisfying $d_\Gamma(p_i, p_j) = d_\Gamma(p_i, m) + d_\Gamma(m, p_j)$ for $1 \leq i < j \leq 3$. A *median graph* is a graph in which every triple of vertices has a *unique* median. Trees and their products are median graphs. See Figure 1 for illustration of the median concept.

Theorem 1.2 ([12]). *If Γ is a median graph, then **0-Ext** $[\Gamma]$ is solvable in polynomial time.*

Karzanov [32] introduced the following LP-relaxation of $\mathbf{0-Ext}[\Gamma]$.

Ext $[\Gamma]$: Given $V \supseteq V_\Gamma$ and $c : \binom{V}{2} \rightarrow \mathbf{Q}_+$,

$$\text{minimize } \sum_{xy \in \binom{V}{2}} c(xy)d(x, y) \text{ over all extensions } (V, d) \text{ of } (V_\Gamma, d_\Gamma).$$

This relaxation $\mathbf{Ext}[\Gamma]$ is a linear program with size polynomial in the input size. Therefore, if for every input (V, c) , $\mathbf{Ext}[\Gamma]$ has an optimal solution that is a 0-extension, then $\mathbf{0-Ext}[\Gamma]$ is solvable in polynomial time. In this case we say that $\mathbf{Ext}[\Gamma]$ is *exact*. In the same paper, Karzanov gave a combinatorial characterization of graphs Γ for which $\mathbf{Ext}[\Gamma]$ is exact. A graph Γ is called a *frame* if

- (1.2) (1) Γ is bipartite,
- (2) Γ has no isometric cycle of length greater than 4, and
- (3) Γ has an orientation o with the property that for every 4-cycle $uv, vv', v'u', u'u$, one has $u \not\prec_o v$ if and only if $u' \not\prec_o v'$.

Here an *isometric cycle* in Γ means a cycle C such that every pair of vertices in C has a shortest path for Γ in this cycle C , and $p \not\prec_o q$ means that edge pq is oriented from q to p by o .

Theorem 1.3 ([32]). $\mathbf{Ext}[\Gamma]$ is exact if and only if Γ is a frame.

Theorem 1.4 ([32]). If Γ is a frame, then $\mathbf{0-Ext}[\Gamma]$ is solvable in polynomial time.

It is noted that the class of frames is *not* closed under taking Cartesian products, whereas the tractability of graphs is preserved under taking Cartesian products. Also it should be noted that $\mathbf{Ext}[\Gamma]$ is the LP-dual to the d_Γ -weighted maximum *multiflow* problem, and $\mathbf{0-Ext}[\Gamma]$ describes a combinatorial dual problem [32, 33]; see also [21, 22, 24, 23] for further elaboration of this duality.

Karzanov [32] also proved the following hardness result. For an undirected graph Γ , an orientation with the property (1.2) (3) is said to be *admissible*. Γ is said to be *orientable* if it has an admissible orientation. Γ is said to be *modular* if every triple of vertices has a (not necessarily unique) median.

Theorem 1.5 ([32]). If Γ is not orientable or not modular, then $\mathbf{0-Ext}[\Gamma]$ is NP-hard.

In fact, a frame is precisely an orientable modular graph with the hereditary property that every isometric subgraph is modular; see [2]. A median graph is an orientable modular graph but the converse is not true. Moreover, a median graph is not necessarily a frame, and a frame is not necessarily a median graph. In [34], Karzanov proved a tractability theorem extending Theorem 1.2. He conjectured that $\mathbf{0-Ext}[\Gamma]$ is tractable for a certain *proper* subclass of orientable modular graphs including frames and median graphs. He also conjectured that $\mathbf{0-Ext}[\Gamma]$ is NP-hard for any graph Γ not in this class.

The main result of this paper is the tractability theorem for *all* orientable modular graphs. Thus the class of tractable graphs is larger than his expectation.

Theorem 1.6. If Γ is orientable modular, then $\mathbf{0-Ext}[\Gamma]$ is solvable in polynomial time.

Combining this result with Theorem 1.5, we obtain a complete classification of the graphs Γ for which $\mathbf{0-Ext}[\Gamma]$ is solvable in polynomial time.

Overview. In proving Theorem 1.6, we employ an axiomatic approach to optimization in orientable modular graphs. This approach is inspired by the theory of discrete convex analysis developed by Murota and his collaborators (including Fujishige, Shioura, and Tamura); see [17, 45, 48, 49, 47] and also [16, Chapter VII]. Discrete convex analysis is a theory of convex functions on integer lattice \mathbf{Z}^n , with the goal of providing a unified framework for polynomially solvable combinatorial optimization problems including network flows, matroids, and submodular functions. The theory that we are going to develop here is, in a sense, *a theory of discrete convex functions on orientable modular graphs*, with the goal of providing a unified framework for polynomially solvable 0-extension problems and related multiflow problems. We believe that our theory establishes a new link between previously unrelated fields, broadens the scope of discrete convex analysis, and opens a new perspective and new research directions.

Let us start with a simple observation to illustrate our basic idea. Consider a path P_m of length m , and consider $\mathbf{0}\text{-Ext}[P_m]$, where P_m is trivially an orientable modular graph. Then $\mathbf{0}\text{-Ext}[P_m]$ for input V, c can be regarded as an optimization problem on the integer lattice \mathbf{Z}^n as follows. Suppose that $V_{P_m} = \{1, 2, 3, \dots, m\}$, and s and $s+1$ are adjacent for $s = 1, 2, \dots, m-1$. Then $d_{P_m}(s, t) = |s - t|$, and $\mathbf{0}\text{-Ext}[P_m]$ is equivalent to the minimization of the function

$$(1.3) \quad \sum_{1 \leq s \leq m} \sum_{1 < j < n} c(sj)|s - \rho_j| + \sum_{1 \leq i < j \leq n} c(ij)|\rho_i - \rho_j|$$

over all $(\rho_1, \rho_2, \dots, \rho_n) \in [0, m]^n \cap \mathbf{Z}^n$. This function is a simple instance of L^\natural -convex functions, one of the fundamental classes of discrete convex functions. We do not give a formal definition of L^\natural -convex functions here. The only important facts for us are the following properties of L^\natural -convex functions in optimization:

- (a) Local optimality implies global optimality.
- (b) The local optimality can be checked by *submodular function minimization*.
- (c) An efficient descent algorithm can be designed based on successive application of submodular function minimization.

As is well-known, submodular functions can be minimized in polynomial time [20, 30, 54]. Actually the function (1.3) can be minimized by successive application of minimum-cut computation [37, 51], a special case of submodular function minimization.

Motivated by this observation, we regard $\mathbf{0}\text{-Ext}[\Gamma]$ as a minimization of a function defined on the vertex set of a product of Γ , which is also orientable modular. We will introduce a class of functions, called *L-convex functions*, on an orientable modular graph. We show that our L-convex function satisfies analogues of (a), (b) and (c) above, and also that a multifacility location function, the objective function of $\mathbf{0}\text{-Ext}[\Gamma]$, is an L-convex function, in our sense, on the product of Γ . Theorem 1.6 is a consequence of these properties.

Let us briefly mention how to define L-convex functions, which constitutes the main body of this paper. Our definition is based on the *Lovász extension* [44], a well-known concept in submodular function theory [16], and a kind of construction of polyhedral complexes, due to Karzanov [32] and Chepoi [13], from a class of modular graphs. Let Γ be an orientable modular graph with admissible orientation o . We call a pair (Γ, o) a *modular complex*. It turns out that (Γ, o) can be viewed as a structure glued together from modular lattices, and gives rise to a simplicial complex as follows. Consider a cube subgraph B of Γ . The digraph \vec{B} oriented by o coincides with the Hasse diagram of

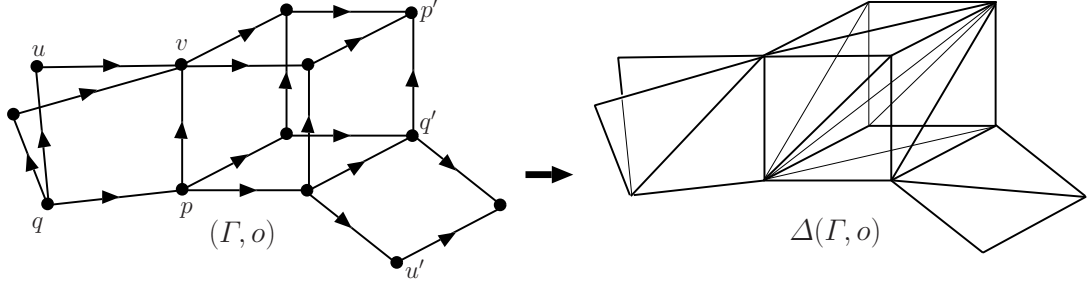


Figure 2: A construction of $\Delta(\Gamma, o)$

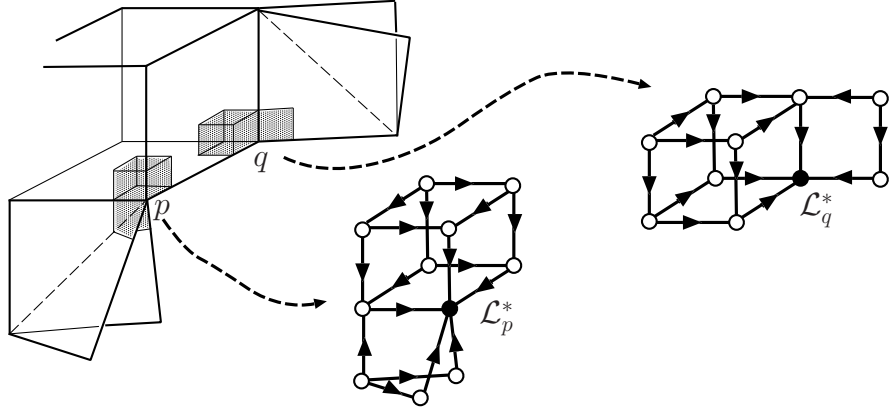


Figure 3: Neighborhood semilattices

a Boolean lattice. Consider the simplicial complex $\Delta(\Gamma, o)$ whose simplices are sets of vertices forming a chain of the Boolean lattice corresponding to some cube subgraph of Γ ; see Figure 2. Each (abstract) simplex is naturally regarded as a simplex in the Euclidean space. $\Delta(\Gamma, o)$ is naturally regarded as a metrized simplicial complex. Then any function $g : V_\Gamma \rightarrow \mathbf{R}$ is extended to $\bar{g} : \Delta(\Gamma, o) \rightarrow \mathbf{R}$ by interpolating g on each simplex linearly; this is an analogue of the Lovász extension. The simplicial complex $\Delta(\Gamma, o)$ enables us to consider the *neighborhood* \mathcal{L}_p^* around each vertex $p \in V_\Gamma$, as well as the local behavior of \bar{g} in \mathcal{L}_p^* . As in Figure 3, neighborhood \mathcal{L}_p^* can be described as a partially ordered set with the unique minimal element p . Then, by restricting \bar{g} to \mathcal{L}_p^* , we obtain a function on \mathcal{L}_p^* associated with each vertex p . In fact, the poset \mathcal{L}_p^* is a *modular semilattice*, a semilattice analogue of a modular lattice introduced by Bandelt, van de Vel, and Verheul [5]. We first define *submodular functions on modular semilattices*, and next define *L-convex functions on modular complex* (Γ, o) as functions g on V_Γ such that \bar{g} is submodular on neighborhood semilattice \mathcal{L}_p^* for each vertex p .

Then the multifacility location function, the objective of $\mathbf{0-Ext}[\Gamma]$ (see (1.1)), is indeed an L-convex function on the n -fold product of Γ , and the optimal solution of $\mathbf{0-Ext}[\Gamma]$ can be obtained by successive application of submodular function minimization on the product of n modular semilattices. Thus our problem reduces to the problem of minimizing submodular function f on the product of modular semilattices $\mathcal{L}_1, \mathcal{L}_2, \dots, \mathcal{L}_n$, where the input of the problem is $\mathcal{L}_1, \mathcal{L}_2, \dots, \mathcal{L}_n$, and an evaluating oracle of f . We do not know whether this problem in general is tractable in the oracle model, but the submodular functions arising from $\mathbf{0-Ext}[\Gamma]$ take a special form; they are *the sum of*

submodular functions with arity 2. Here the *arity* of a function f is the number of variables of f . Namely, if a function f on $\mathcal{L} = \mathcal{L}_1 \times \mathcal{L}_2 \times \cdots \times \mathcal{L}_n$ is represented as

$$f(x) = h(x_{i_1}, x_{i_2}, \dots, x_{i_k}) \quad (x = (x_1, x_2, \dots, x_n) \in \mathcal{L})$$

for some function h on $\mathcal{L}_{i_1} \times \mathcal{L}_{i_2} \times \cdots \times \mathcal{L}_{i_k}$ with $i_1 < i_2 < \cdots < i_k$, then the arity of f is (at most) k . See (1.1); our objective function is a weighted sum of distance functions, which have arity 2. This type of optimization problem with bounded arity is well-studied in the literature of *valued CSP* (*valued constraint satisfaction problem*) [7, 42, 53, 59]. Valued CSP deals with minimization of a sum of functions f_i ($i = 1, 2, \dots, m$), where the arity k_i of each f_i is a part of the input; namely the input consists of all values of all functions f_i . Valued CSP admits an integer programming formulation, and its natural LP relaxation is called the *basic LP-relaxation*. Recently, Thapper and Živný [56] discovered a surprising criterion for the basic LP-relaxation of valued CSP to exactly solve the original valued CSP instance. They proved that if the class of valued CSP (the class of input objective functions) has a certain nice *fractional polymorphism* (a certain set of linear inequalities which any input function satisfies), then the basic LP-relaxation is exact. We prove that the class of submodular functions on modular semilattice admits such a fractional polymorphism. Then the sum of submodular functions with bounded arity can be minimized in polynomial time. Consequently we can solve **0-Ext**[Γ] in polynomial time.

We believe that our classes of functions deserve to be called submodular and L-convex. Indeed, they include not only (ordinary) submodular/L-convex functions but also other submodular/L-convex-type functions. Examples are *bisubmodular functions* [9, 50, 52] (see [16, Section 3.5]), *multimatroid rank functions* by Bouchet [8], *submodular functions on trees* by Kolmogorov [38], *k-submodular functions* by Huber and Kolmogorov [27] (also see [18]), and *skew-bisubmodular functions* by Huber, Krokshin and Powell [29] (also see [19, 29, 28]). Moreover, combinatorial dual problems arising from a large class of (well-behaved) multicommodity flow problems, discussed in [21, 22, 24, 23, 31, 32, 33], fall into submodular/L-convex function minimization in our sense. This can be understood as a multiflow analogue of a fundamental fact in network flow theory: the minimum cut problem, the dual of maxflow problem, is a submodular function minimization. The detailed discussion on these topics will be given in a separate paper [26]; some of the results were announced by [25].

Organization. In Section 2, we first explain basic notions of valued CSP and the Thapper-Živný criterion (Theorem 2.1) on the exactness of the basic LP relaxation. We then describe basic facts on modular graphs and modular lattices. In Section 3, we develop a theory of submodular functions on modular semilattices. We show that our submodular function satisfies the Thapper-Živný criterion, and that a sum of submodular functions with bounded arity can be minimized in polynomial time. In Section 4, we first explore several structural properties of orientable modular graphs. Based on the above mentioned idea, we define L-convex functions, and prove that our L-convex functions indeed have properties analogous to (a), (b) and (c) above. In Section 5, we formulate **0-Ext**[Γ] as an optimization problem on a modular complex. We show that a multifacility location function, the objective function of **0-Ext**[Γ], is indeed an L-convex function, and we prove Theorem 1.6. Our framework is applicable to a certain weighted version of **0-Ext**[Γ]. As a corollary, we give a generalization of Theorem 1.6 to general metrics, which completes classification of metrics μ for which the 0-extension problem on μ is polynomial time solvable (Theorem 5.9). In the last section (Section 6), we discuss a connection to a dichotomy theorem of finite-valued CSP obtained by Thapper and

Živný [57] after the first submission of this paper. In fact, the complexity dichotomy (of form “either P or NP-hard”) of **0-Ext**, established in this paper, can be viewed as a special case of their dichotomy theorem of finite-valued CSP.

Notation. Let \mathbf{Z}, \mathbf{Q} , and \mathbf{R} denote the sets of integers, rationals, and reals, respectively. Let $\overline{\mathbf{R}} := \mathbf{R} \cup \{\infty\}$ and $\overline{\mathbf{Q}} := \mathbf{Q} \cup \{\infty\}$, where ∞ is an infinity element and is treated as: $\infty \cdot 0 = 0$, $x < \infty$ ($x \in \mathbf{R}$), $\infty + x = \infty$ ($x \in \overline{\mathbf{R}}$), $x \cdot \infty = \infty$ ($a \in \mathbf{R} : a > 0$). Let $\mathbf{Z}_+, \mathbf{Q}_+$, and \mathbf{R}_+ denote the sets of nonnegative integers, nonnegative rationals, and nonnegative reals, respectively. For a function $f : X \rightarrow \overline{\mathbf{R}}$ on a set X , let $\text{dom } f$ denote the set of elements $x \in X$ with $f(x) \neq \infty$.

For a graph Γ , the vertex set and the edge set are denoted by V_Γ and E_Γ , respectively. For a vertex subset X , $\Gamma[X]$ denotes the subgraph of Γ induced by X . For a nonnegative edge-length $h : E_\Gamma \rightarrow \mathbf{R}_+$, $d_{\Gamma,h}$ denotes the shortest path metric on V_Γ with respect to the edge-length h . When $h(e) = 1$ for every edge e , $d_{\Gamma,h}$ is denoted by d_Γ . A path is represented by a chain (p_1, p_2, \dots, p_n) of vertices with $p_i p_{i+1} \in E_\Gamma$. The *Cartesian product* $\Gamma \times \Gamma'$ of graphs Γ and Γ' is the graph with vertex set $V_\Gamma \times V_{\Gamma'}$ and edge set given as: (p, p') and (q, q') are connected by an edge if and only if $p = q$ and $p'q' \in E_{\Gamma'}$ or $p' = q'$ and $pq \in E_\Gamma$. The n -fold Cartesian product $\Gamma \times \Gamma \times \dots \times \Gamma$ of Γ is denoted by Γ^n . In this paper, graphs and posets (partially ordered sets) are supposed to be finite.

2 Preliminaries

In this section, we give preliminary arguments for valued CSP, and modular graphs and modular (semi)lattices. Our references are [41, 59] for valued CSP and [3, 5, 6, 13, 58] for modular graphs and lattices. A further discussion on valued CSP is given in Section 6.

2.1 Valued CSP and fractional polymorphism

Let D_1, D_2, \dots, D_n be finite sets, and let $D := D_1 \times D_2 \times \dots \times D_n$. A *constraint* on D is a function $f : D_{i_1} \times D_{i_2} \times \dots \times D_{i_k} \rightarrow \overline{\mathbf{R}}$ for some $i_1 < i_2 < \dots < i_k$, where $I_f := \{i_1, i_2, \dots, i_k\}$ is called the *scope* of f , and $k_f := k$ is called the *arity* of f . Let $D_{I_f} := D_{i_1} \times D_{i_2} \times \dots \times D_{i_k}$, and for $x = (x_1, x_2, \dots, x_n) \in D$, let $x_{I_f} := (x_{i_1}, x_{i_2}, \dots, x_{i_k}) \in D_{I_f}$. The *valued CSP* (*valued constraint satisfaction problem*) is:

VCSP: Given a set \mathcal{F} of constraints on D ,

$$\text{minimize } \sum_{f \in \mathcal{F}} f(x_{I_f}) \text{ over all } x = (x_1, x_2, \dots, x_n) \in D.$$

The input of VCSP is the set of all values of all constraints in \mathcal{F} , and hence its size is estimated by $O(|\mathcal{F}|N^K B)$, where $N := \max_{1 \leq i \leq n} |D_i|$, $K := \max_{f \in \mathcal{F}} k_f$, and B is the bit size to represent constraints in \mathcal{F} . By a *constraint language* we mean a (possibly infinite) set Λ of constraints. A constraint in Λ is called a Λ -*constraint*. Let **VCSP** $[\Lambda]$ denote the subclass of **VCSP** such that the input is restricted to a set of Λ -constraints.

The minimum 0-extension problem **0-Ext** $[\Gamma]$ is formulated as an instance of **VCSP**. Let $D_i := V_\Gamma$ for $i = 1, 2, \dots, n$. Define constraints $g_i : D_i \rightarrow \mathbf{R}$ and $f_{ij} : D_i \times D_j \rightarrow \mathbf{R}$ by

$$\begin{aligned} (2.1) \quad g_i(\rho_i) &:= \sum_{s \in V_\Gamma} c(si) d_\Gamma(s, \rho_i) \quad (\rho_i \in D_i), \\ f_{ij}(\rho_i, \rho_j) &:= c(ij) d_\Gamma(\rho_i, \rho_j) \quad ((\rho_i, \rho_j) \in D_i \times D_j). \end{aligned}$$

Define the input \mathcal{F} of **VCSP** by

$$\mathcal{F} := \{g_i \mid 1 \leq i \leq n\} \cup \{f_{ij} \mid 1 \leq i < j \leq n\}.$$

Notice that the size of \mathcal{F} is polynomial in n , Γ , and the bit size representing c . Hence **0-Ext** $[\Gamma]$ is a particular subclass of **VCSP**.

VCSP admits the following integer programming formulation:

$$(2.2) \quad \begin{aligned} \text{Min.} \quad & \sum_{f \in \mathcal{F}} \sum_{y \in \text{dom } f} f(y) \lambda_{f,y} \\ \text{s.t.} \quad & \sum_{y \in \text{dom } f: y_i = a} \lambda_{f,y} = \mu_{i,a} \quad (f \in \mathcal{F}, i \in I_f, a \in D_i) \\ & \sum_{a \in D_i} \mu_{i,a} = 1 \quad (1 \leq i \leq n), \\ & \lambda_{f,y} \in \{0, 1\} \quad (f \in \mathcal{F}, y \in \text{dom } f), \\ & \mu_{i,a} \in \{0, 1\} \quad (1 \leq i \leq n, a \in D_i). \end{aligned}$$

Indeed, for each i there uniquely exists $a_i \in D_i$ with $\mu_{i,a_i} = 1$. Also for $f \in \mathcal{F}$ there uniquely exists $y \in \text{dom } f$ such that $\lambda_{f,y} = 1$ and $y_i = a_i$ for $i \in I_f$. Define $x = (x_1, x_2, \dots, x_n)$ by $x_i := a_i$. Then $\lambda_{f,y} = 1$ if and only if $x_{I_f} = y$. Therefore we obtain a solution $x = (x_1, x_2, \dots, x_n)$ of **VCSP** with the same objective value. Conversely, for a solution $x = (x_1, x_2, \dots, x_n)$ of **VCSP**, define $\mu_{i,a} := 1$ if $x_i = a$, and $\lambda_{f,y} := 1$ if $x_{I_f} = y$. The other variables are defined as zero. Then we obtain a solution of (2.2) with the same objective value.

Observe that there are $O(|\mathcal{F}|N^K + nN)$ variables and $O(|\mathcal{F}|KN + n)$ constraints. Therefore the size of this IP is bounded by a polynomial of the input size. The *basic LP relaxation* (BLP) is the linear problem obtained by relaxing the 0-1 constraints $\lambda_{f,y} \in \{0, 1\}$ and $\mu_{i,a} \in \{0, 1\}$ into $\lambda_{f,y} \geq 0$ and $\mu_{i,a} \geq 0$, respectively. In particular BLP can be solved in (strongly) polynomial time.

Recently Thapper and Živný [56] discovered a surprisingly powerful criterion for which BLP solves **VCSP**. To describe their result, let us introduce some notions. For a constraint language Λ , BLP is said to be *exact* for Λ if for every input $\mathcal{F} \subseteq \Lambda$, the optimal value of BLP coincides with the optimal value of **VCSP** $[\Lambda]$. An *operation* on D_i is a function $D_i \times D_i \rightarrow D_i$. A (separable) operation ϑ on D is a function $D \times D \rightarrow D$ such that ϑ is represented as

$$\vartheta(x, y) = (\vartheta_1(x_1, y_1), \vartheta_2(x_2, y_2), \dots, \vartheta_n(x_n, y_n)) \quad (x, y \in D)$$

for some operations ϑ_i on D_i for $i = 1, 2, \dots, n$. A *fractional operation* ω is a function from the set of all operations to \mathbf{R}_+ such that the total sum $\sum \omega(\vartheta)$ over all operations ϑ is 1. We denote a fractional operation ω by the form of a *formal convex combination* $\sum \omega(\vartheta) \vartheta$ of operations ϑ . The *support* of ω is the set of operations ϑ with $\omega(\vartheta) > 0$. For a constraint language Λ , a *fractional polymorphism* is a fractional operation $\sum \omega(\vartheta) \vartheta$ on D such that it satisfies

$$(2.3) \quad \frac{f(x) + f(y)}{2} \geq \sum_{\vartheta} \omega(\vartheta) f(\vartheta(x, y)) \quad (f \in \Lambda, x, y \in D_{I_f}),$$

where ϑ is regarded an operation on D_{I_f} by $(\vartheta(x, y))_i := \vartheta_i(x_i, y_i)$ for $i \in I_f$. For example, if D_i is a lattice for each i , then $\frac{1}{2} \wedge + \frac{1}{2} \vee$ is nothing but a fractional polymorphism for submodular functions, i.e., functions f satisfying $f(p) + f(q) \geq f(p \wedge q) + f(p \vee q)$ for $p, q \in D$.

Theorem 2.1 (Special case of [56, Theorem 5.1]). *If a constraint language Λ admits a fractional polymorphism ω such that the support of ω contains a semilattice operation, then BLP is exact for Λ , and hence $\mathbf{VCSP}[\Lambda]$ can be solved in polynomial time.*

Here a *semilattice operation* is an operation ϑ satisfying $\vartheta(a, a) = a$, $\vartheta(a, b) = \vartheta(b, a)$, and $\vartheta(\vartheta(a, b), c) = \vartheta(a, \vartheta(b, c))$ for $a, b, c \in D$. Although the feasible region of BLP is not necessarily an integral polytope, we can check whether there exists an optimal solution x with $x_i = a \in D_i$ by comparing the optimal values of BLP for the input \mathcal{F} and for $\mathcal{F}_{i,a}$, which is the set of cost functions obtained by fixing variable x_i to a for each cost function on \mathcal{F} . Necessarily BLP is exact for $\mathcal{F}_{i,a}$ if there is an optimal solution x with $x_i = a$. Hence, after n fixing procedures, we obtain an optimal solution x .

Remark 2.2. In the setting in [56], D_i is the same set \tilde{D} for all i . Our setting reduces to this case by taking the disjoint union of D_i as \tilde{D} , and extending each cost function $f : D_{I_f} \rightarrow \overline{\mathbf{R}}$ to $f : \tilde{D}^{k_f} \rightarrow \overline{\mathbf{R}}$ by $f(x_{i_1}, x_{i_2}, \dots, x_{i_k}) := \infty$ for $(x_{i_1}, x_{i_2}, \dots, x_{i_k}) \notin D_{i_1} \times D_{i_2} \times \dots \times D_{i_k}$. Without such a reduction, their proof also works for our setting in a straightforward way.

2.2 Modular metric spaces and modular graphs

For a metric space (X, d) , the (*metric*) *interval* $I(x, y)$ of $x, y \in X$ is defined as

$$I(x, y) := \{z \in X \mid d(x, z) + d(z, y) = d(x, y)\}.$$

For two subsets A, B , $d(A, B)$ denotes the infimum of distances between A and B , i.e.,

$$d(A, B) = \inf_{x \in A, y \in B} d(x, y).$$

For $x_1, x_2, x_3 \in X$, an element m in $I(x_1, x_2) \cap I(x_2, x_3) \cap I(x_3, x_1)$ is called a *median* of x_1, x_2 , and x_3 . A metric space (X, d) is said to be *modular* if every triple of elements in X has a median. In particular, a graph Γ is modular if and only if the shortest path metric space (V_Γ, d_Γ) is modular. We will often use the following characterization of modular graphs.

Lemma 2.3 ([5, Proposition 1.7]; see [58, Proposition 6.2.6, Chapter I]). *A connected graph Γ is modular if and only if*

- (1) Γ is bipartite, and
- (2) for vertices p, q and neighbors p_1, p_2 of p with $d_\Gamma(p, q) = 1 + d_\Gamma(p_1, q) = 1 + d_\Gamma(p_2, q)$, there exists a common neighbor p^* of p_1, p_2 with $d_\Gamma(p, q) = 2 + d_\Gamma(p^*, q)$.

The condition (2) is called the *quadrangle condition* [3, 13] (or the *semimodularity condition* in [5, 58]).

Lemma 2.4. *For a modular graph, every admissible orientation is acyclic.*

Proof. Suppose indirectly that the statement is false. Take a vertex p belonging to a directed cycle, and take a directed cycle C containing p with $\sum_{u \in V_C} d_\Gamma(p, u)$ minimum. The length k of C is at least four (by simpleness and bipartiteness). By the definition of admissible orientation, $k = 4$ is impossible. Hence $k > 4$. Take a vertex q in C with $d_\Gamma(p, q)$ maximum. Take two neighbors q', q'' of q in C . Then $d_\Gamma(q, p) = d_\Gamma(q', p) + 1 = d_\Gamma(q'', p) + 1$ (by the maximality of q and the bipartiteness of Γ). By the quadrangle condition, there is a common neighbor q^* of q', q'' with $d_\Gamma(q, p) = d_\Gamma(p, q^*) + 2$. Here the cycle C' obtained from C by replacing q by q^* is a directed cycle, since the orientation is admissible. Then we have $\sum_{u' \in V_{C'}} d_\Gamma(p, u') < \sum_{u \in V_C} d_\Gamma(p, u)$. This contradicts the minimality of C . \square

2.2.1 Orbits and orbit-invariant functions

Let Γ be a modular graph. Edges e and e' are said to be *projective* if there is a sequence $(e = e_0, e_1, e_2, \dots, e_m = e')$ of edges such that e_i and e_{i+1} belong to a common 4-cycle and share no common vertex. We will use the following criterion for two edges to belong to a common orbit.

Lemma 2.5. *Let Γ be a modular graph. For edges pq and $p'q'$, suppose that $d_\Gamma(p, p') = d_\Gamma(q, q')$ and $d_\Gamma(p, q') = d_\Gamma(p, p') + 1 = d_\Gamma(p', q)$.*

- (1) *pq and $p'q'$ are projective.*
- (2) *In addition, if Γ has an admissible orientation o , then $p \searrow_o q$ implies $p' \searrow_o q'$.*

Proof. We use the induction on $k := d_\Gamma(p, p') = d_\Gamma(q, q')$. The case of $k = 1$ is obvious. Take a neighbor p^* of p with $d_\Gamma(p^*, p') = d_\Gamma(p, p') - 1 = k - 1$. Then $d_\Gamma(p^*, q') = k$. By the quadrangle condition for p, q, p^*, q' , there is a common neighbor q^* of q, p^* with $d_\Gamma(q^*, q') = k - 1$. Also $d_\Gamma(q^*, p') = k$. Obviously pq and p^*q^* are projective, and $p \searrow_o q$ implies $p^* \searrow_o q^*$. Apply the induction for p^*q^* and $p'q'$. \square

An *orbit* is an equivalence class of the projectivity relation. The (disjoint) union of several orbits is called an *orbit-union*. For an orbit-union U , Γ/U is the graph obtained by contracting all edges *not* in U and by identifying multiple edges. The vertex in Γ/U corresponding to $p \in V_\Gamma$ is denoted by p/U . The graph Γ/U is also modular, and any shortest path in Γ induces a shortest path in Γ/U as follows.

Lemma 2.6 ([1], also see [34]). *Let Γ be a modular graph, and U an orbit-union.*

- (1) *Γ/U is a modular graph.*
- (2) *For every $p, q \in V_\Gamma$, every shortest (p, q) -path P , and every (p, q) -path P' , we have $|P \cap U| \leq |P' \cap U|$.*
- (3) *For every $p, q \in V_\Gamma$ and every shortest (p, q) -path P , the image P/U of P is a shortest $(p/U, q/U)$ -path in Γ/U .*

In particular, for any partition \mathcal{U} of E_Γ into orbit-unions, we have

$$d_\Gamma(p, q) = \sum_{U \in \mathcal{U}} d_{\Gamma/U}(p/U, q/U) = \sum_{Q: \text{orbit}} d_{\Gamma/Q}(p/Q, q/Q) \quad (p, q \in V_\Gamma).$$

A function h on edge set E_Γ is called *orbit-invariant* if $h(e) = h(e')$ provided e and e' belong to the same orbit. For an orbit Q , let h_Q denote the value of h on Q . An orbit-invariant function h is said to be *nonnegative* if $h(e) \geq 0$ for $e \in E_\Gamma$, and is said to be *positive* if $h(e) > 0$ for $e \in E_\Gamma$. For a constant $c \geq 0$, if $h(e) = c$ for all edges e , then h is simply denoted by c ; in particular $d_\Gamma = d_{\Gamma, 1}$. By taking the value of h of the preimage, we can define a function on the edge set of Γ/U for any orbit-union U , which is also orbit-invariant in Γ/U and is denoted by h . By Lemma 2.6 (2), the shortest path structures of (V_Γ, d_Γ) and $(V_\Gamma, d_{\Gamma, h})$ are the same in the following sense:

Lemma 2.7. *If an orbit-invariant function h is nonnegative, then (1) implies (2), where*

- (1) *P is a shortest (p, q) -path with respect to 1,*
- (2) *P is a shortest (p, q) -path with respect to h .*

If h is positive, then the converse also holds.

As a consequence of Lemmas 2.6 and 2.7, for any partition \mathcal{U} of E_Γ into orbit-unions, we have

$$(2.4) \quad d_{\Gamma,h}(p, q) = \sum_{U \in \mathcal{U}} d_{\Gamma/U,h}(p/U, q/U) = \sum_{Q: \text{orbit}} h_Q d_{\Gamma/Q,1}(p/Q, q/Q).$$

2.2.2 Convex sets and gated sets

Let (X, d) be a metric space. A subset $Y \subseteq X$ is called *convex* if $I(p, q) \subseteq Y$ for every $p, q \in Y$. A subset $Y \subseteq X$ is called *gated* if for every $p \in X$ there is $p^* \in Y$, called a *gate* of p at Y , such that $d(p, q) = d(p, p^*) + d(p^*, q)$ holds for every $q \in Y$. One can easily see that gate p^* is uniquely determined for each p [15, p. 112]. Therefore we obtain a map $\text{Pr}_Y : X \rightarrow Y$ by defining $\text{Pr}_Y(p)$ to be the gate of p at Y .

Theorem 2.8 ([15]). *Let A and A' be gated subsets of (X, d) and let $B := \text{Pr}_A(A')$ and $B' := \text{Pr}_{A'}(A)$.*

- (1) Pr_A and $\text{Pr}_{A'}$ induce isometries, inverse to each other, between B' and B .
- (2) For $p \in A$ and $p' \in A'$, the following conditions are equivalent:
 - (i) $d(p, p') = d(A, A')$.
 - (ii) $p = \text{Pr}_A(p')$ and $p' = \text{Pr}_{A'}(p)$.
- (3) B and B' are gated, and $\text{Pr}_B = \text{Pr}_A \circ \text{Pr}_{A'}$ and $\text{Pr}_{B'} = \text{Pr}_{A'} \circ \text{Pr}_A$.

As remarked in [15], every gated set is convex (see the proof of Lemma 2.9 below). The converse is not true in general, but is true for modular graphs. The following useful characterization of convex (gated) sets in a modular graph is due to Chepoi [11]. Here, for a graph Γ , a subset Y of vertices is said to be convex (resp. gated) if Y is convex (resp. gated) in (V_Γ, d_Γ) .

Lemma 2.9 ([11]). *Let Γ be a modular graph. For $Y \subseteq V_\Gamma$, the following conditions are equivalent:*

- (1) Y is convex.
- (2) Y is gated.
- (3) $\Gamma[Y]$ is connected and $I(p, q) \subseteq Y$ holds for every $p, q \in Y$ with $d_\Gamma(p, q) = 2$.

We give a proof for the convenience of readers as the original paper is in Russian.

Proof. d_Γ is denoted by d . (1) \Rightarrow (3) is obvious. We show (3) \Rightarrow (1). Take $p, q \in Y$, and take $a \in I(p, q)$. We are going to show $a \in Y$. Since $\Gamma[Y]$ is connected, we can take a path $P = (p = p_0, p_1, \dots, p_k = q)$ with $p_i \in Y$. Take such a path P with $\kappa_P := \sum_{i=0}^k d(a, p_i)$ minimum. If $d(a, p_{i-1}) < d(a, p_i) > d(a, p_{i+1})$ for some i , then, by the quadrangle condition in Lemma 2.3, there is a common neighbor p^* of p_{i-1}, p_{i+1} with $d(a, p^*) = d(a, p_i) - 2$. Since $I(p_{i-1}, p_{i+1}) \subseteq Y$ by (3), p^* belongs to Y . Then we can replace p_i by p^* in P to obtain another path P' connecting p, q with $\kappa_{P'} = \kappa_P - 2$; a contradiction to the minimality. Therefore there is no index j with $d(a, p_{j-1}) < d(a, p_j) > d(a, p_{j+1})$. Thus there is a unique index i with $d(a, p_i)$ minimum. Then we have $d(p, p_i) + d(p_i, a) = d(p, a)$ and $d(q, p_i) + d(p_i, a) = d(q, a)$. By $a \in I(p, q)$, we have $d(p, q) = d(p, a) + d(a, q) = d(p, p_i) + d(p_i, q) + 2d(p_i, a) \geq d(p, q) + 2d(p_i, a)$. Hence we must have $d(p_i, a) = 0$, implying $a = p_i \in Y$.

We show (2) \Rightarrow (1). As already mentioned, any gated set is convex. Indeed, suppose that Y is gated. Take $p, q \in Y$, and take $a \in I(p, q)$. Consider the gate a^* of a in Y . Then $d(p, a) = d(p, a^*) + d(a^*, a)$ and $d(q, a) = d(q, a^*) + d(a^*, a)$. Since $a \in I(p, q)$, we have $d(p, q) = d(p, a) + d(a, q) = d(p, a^*) + d(a^*, q) + 2d(a^*, a) \geq d(p, q) + 2d(a^*, a)$, implying $d(a^*, a) = 0$ and $a = a^* \in Y$. Thus we get (2) \Rightarrow (1).

Finally we show (1) \Rightarrow (2). Suppose that Y is convex. Let p be an arbitrary vertex. Let p^* be a point in Y satisfying $d(p, Y) = d(p, p^*)$. We show that p^* is a gate of p at Y . Take arbitrary $q \in Y$. Consider a median m of p, q, p^* . By convexity, m belongs to Y , and also $m \in I(p^*, p)$. By definition of p^* , it must hold $p^* = m$. Thus $d(p, q) = d(p, p^*) + d(p^*, q)$ holds for every $q \in Y$. This means that p^* is the gate of p , and therefore Y is gated. \square

2.3 Modular lattices and modular semilattices

Let \mathcal{L} be a partially ordered set (poset) with partial order \preceq . For $a, b \in \mathcal{L}$, the (unique) minimum common upper bound, if it exists, is denoted by $a \vee b$, and the (unique) maximum common lower bound, if it exists, is denoted by $a \wedge b$. \mathcal{L} is said to be a *lattice* if both $a \vee b$ and $a \wedge b$ exist for every $a, b \in \mathcal{L}$, and said to be a (*meet-*)*semilattice* if $a \wedge b$ exists for every $a, b \in \mathcal{L}$. In a semilattice, if a and b have a common upper bound, then $a \vee b$ exists. Such (a, b) is said to be *bounded*. By the expression “ $a \vee b \in \mathcal{L}$ ” we mean that $a \vee b$ exists. A pair (a, b) is said to be *comparable* if $a \preceq b$ or $b \preceq a$, and *incomparable* otherwise. We say “ b covers a ” if $a \prec b$ and there is no $c \in \mathcal{L}$ with $a \prec c \prec b$, where $a \prec b$ means $a \preceq b$ and $a \neq b$. The maximum element (universal upper bound) and the minimum element (universal lower bound), if they exist, are denoted by $\mathbf{1}$ and $\mathbf{0}$, respectively. For $a \preceq b$, the interval $\{c \in \mathcal{L} \mid a \preceq c \preceq b\}$ is denoted by $[a, b]$. A *chain* from a to b is a sequence $(a = u_0, u_1, u_2, \dots, u_k = b)$ with $u_{i-1} \prec u_i$ for $i = 1, 2, \dots, k$; the number k is the length of the chain. The length $r[a, b]$ of the interval $[a, b]$ is defined as the maximum length of a chain from a to b . The rank $r(a)$ of an element a is defined by $r(a) = r[\mathbf{0}, a]$. An *atom* is an element of rank 1. The *covering graph* of a poset \mathcal{L} is the underlying undirected graph of the Hasse diagram of \mathcal{L} .

A lattice \mathcal{L} is called *modular* if $a \vee (b \wedge c) = (a \vee b) \wedge c$ for every $a, b, c \in \mathcal{L}$ with $a \preceq c$. Modular lattices are also characterized by the modular equality of the rank function.

Lemma 2.10 (See [6, Chapter III, Corollary 1]). *A lattice \mathcal{L} is modular if and only if*

$$r(a) + r(b) = r(a \vee b) + r(a \wedge b) \quad (a, b \in \mathcal{L}).$$

A lattice \mathcal{L} is called *complemented* if for every $p \in \mathcal{L}$ there is an element q , called a *complement* of p , such that $p \vee q = \mathbf{1}$ and $p \wedge q = \mathbf{0}$, and *relatively complemented* if $[a, b]$ is complemented for every $a, b \in \mathcal{L}$ with $a \preceq b$.

Theorem 2.11 (See [6, Chapter IV, Theorem 4.1]). *Let \mathcal{L} be a modular lattice. The following conditions are equivalent:*

- (1) \mathcal{L} is complemented.
- (2) \mathcal{L} is relatively complemented.
- (3) Every element is the join of atoms.
- (4) $\mathbf{1}$ is the join of atoms.

Modular semilattice. The modularity concept has been extended for semilattices by Bandelt, van de Vel, and Verheul [4]. A semilattice \mathcal{L} is said to be *modular* if $[\mathbf{0}, p]$ is a modular lattice for every $p \in \mathcal{L}$, and $a \vee b \vee c \in \mathcal{L}$ provided $a \vee b, b \vee c, c \vee a \in \mathcal{L}$. A modular semilattice is said to be *complemented* if $[\mathbf{0}, p]$ is a complemented modular lattice for every $p \in \mathcal{L}$.

It is known that a lattice is modular if and only if its covering graph is modular; see [58, Proposition 6.2.1]. A modular semilattice is characterized by an analogous property as follows.

Theorem 2.12 ([5, Theorem 5.4]). *A semilattice is modular if and only if its covering graph is modular.*

The Hasse diagram of \mathcal{L} is admissibly oriented since every 4-cycle is a form of $(p, p \wedge q, q, p \vee q)$.

Corollary 2.13. *The covering graph of a modular semilattice is orientable modular.*

Let \mathcal{L} be a modular semilattice and let Γ be the covering graph of \mathcal{L} , which is orientable modular. An immediate consequence of the Jordan-Dedekind chain condition for modular lattices is:

$$(2.5) \quad \text{For } p, q \in \mathcal{L} \text{ with } p \preceq q, \text{ we have } I(p, q) = [p, q] \text{ and } d_\Gamma(p, q) = r[p, q].$$

A (positive) *valuation* of \mathcal{L} is a function on \mathcal{L} satisfying

$$(2.6) \quad v(q) - v(p) > 0 \quad (p, q \in \mathcal{L} : p \prec q),$$

$$(2.7) \quad v(p) + v(q) = v(p \wedge q) + v(p \vee q) \quad (p, q \in \mathcal{L} : \text{bounded}).$$

This is a natural extension of a valuation of a modular lattice; see [6, Chapter III, 50] (we follow the terminology in the third edition of this book). In particular, the rank function r is a valuation. For p, q with $p \preceq q$, let $v[p, q]$ denote $v(q) - v(p)$. Valuations and orbit-invariant functions are related in the following way.

Lemma 2.14. (1) *For a valuation v on \mathcal{L} , the edge-length h on Γ defined by*

$$h(pq) := v(q) - v(p) \quad (p, q \in \mathcal{L} : q \text{ covers } p)$$

is a positive orbit-invariant function, and satisfies

$$v[p, q] = d_{\Gamma, h}(p, q) \quad (p, q \in \mathcal{L} : p \preceq q).$$

(2) *For a positive orbit-invariant function h on Γ , a function v on \mathcal{L} defined by*

$$v(p) := d_{\Gamma, h}(\mathbf{0}, p) \quad (p \in \mathcal{L})$$

is a valuation.

Proof. (1). The positivity of h follows from (2.6). The orbit invariance of h follows from (2.7) and the observation that every 4-cycle of Γ is the form of $(p, p \wedge q, q, p \vee q)$, where $p \vee q$ covers p and q , and $p \wedge q$ is covered by p and q . We show the latter part by induction on $r[p, q]$; the case $r[p, q] = 1$ is obvious. Take $p' \in [p, q] = I(p, q)$ such that p' covers p . By induction, we have $d_{\Gamma, h}(p', q) = v[p', q]$. By Lemma 2.7 and (2.5), we have $d_{\Gamma, h}(p, q) = h(pp') + d_{\Gamma, h}(p', q) = v[p, q]$.

(2). By Lemma 2.7 and (2.5), if q covers p , then $v(q) - v(p) = h(pq) > 0$, implying (2.6). For a bounded pair (p, q) , take maximal chains $(p \wedge q = p_0, p_1, \dots, p_k = p)$ and

$(p \wedge q = q_0, q_1, \dots, q_l = q)$. Let $a_{i,j} := p_i \vee q_j$. By modularity, we see that $a_{i+1,j+1}$ covers $a_{i+1,j}$ and $a_{i,j+1}$, and $a_{i,j}$ is covered by $a_{i+1,j}$ and $a_{i,j+1}$; in particular $a_{i+1,j+1} = a_{i+1,j} \vee a_{i,j+1}$ and $a_{i,j} = a_{i+1,j} \wedge a_{i,j+1}$. Therefore $v(p) + v(q) - v(p \wedge q) - v(p \vee q) = \sum_{i,j} (v(a_{i+1,j}) + v(a_{i,j+1}) - v(a_{i+1,j+1}) - v(a_{i,j})) = \sum_{i,j} (h(a_{i+1,j}a_{i,j}) - h(a_{i,j+1}a_{i+1,j+1}))$. By the orbit invariance of h , all summands are zero, implying (2.7). \square

Consider the case where \mathcal{L} is the product $\mathcal{L}_1 \times \mathcal{L}_2$ of two modular semilattices $\mathcal{L}_1, \mathcal{L}_2$. For a valuation v on \mathcal{L} , define $v_i : \mathcal{L}_i \rightarrow \mathbf{R}$ ($i = 1, 2$) by

$$(2.8) \quad v_1(p_1) := v(p_1, \mathbf{0}) \quad (p_1 \in \mathcal{L}_1), \quad v_2(p_2) := v(\mathbf{0}, p_2) \quad (p_2 \in \mathcal{L}_2).$$

Then v_i is a valuation on \mathcal{L}_i for $i = 1, 2$, and satisfies

$$(2.9) \quad v(p) = v_1(p_1) + v_2(p_2) - v(\mathbf{0}) \quad (p = (p_1, p_2) \in \mathcal{L}).$$

Conversely, for a valuation v_i on \mathcal{L}_i ($i = 1, 2$), define $v : \mathcal{L} \rightarrow \mathbf{R}$ by

$$(2.10) \quad v(p) := v_1(p_1) + v_2(p_2) \quad (p = (p_1, p_2) \in \mathcal{L}).$$

Then v is a valuation on \mathcal{L} .

In the sequel, a modular semilattice \mathcal{L} is supposed to be endowed with some valuation v . If \mathcal{L} is the product of modular semilattices \mathcal{L}_i , then the valuation of each \mathcal{L}_i is defined by (2.8), and is also denoted by v . For modular semilattices \mathcal{L}_1 and \mathcal{L}_2 , the valuation v of $\mathcal{L}_1 \times \mathcal{L}_2$ is defined to be the sum of valuations of \mathcal{L}_1 and \mathcal{L}_2 according to (2.10). Also a modular semilattice \mathcal{L} is regarded as a metric space by the shortest path metric of its covering graph Γ with respect to a positive orbit-invariant function in Lemma 2.14 (1). The corresponding metric function is denoted by $d = d_{\mathcal{L}}$. We give basic properties of metric intervals of \mathcal{L} .

Lemma 2.15. *For $p, q \in \mathcal{L}$, we have the following.*

- (1) $d(p, q) = v[p \wedge q, p] + v[p \wedge q, q]$.
- (2) $I(p, q) = \{a \vee b \mid a \in [p \wedge q, p], b \in [p \wedge q, q] : a \vee b \text{ exists}\}$.
- (3) If $c = a \vee b$ for $a \in [p \wedge q, p], b \in [p \wedge q, q]$, then $a = p \wedge c$ and $b = q \wedge c$.
- (4) For $u, u' \in I(p, q)$, it holds $u \wedge u' = (u \wedge u' \wedge p) \vee (u \wedge u' \wedge q)$; in particular $u \wedge u' \in I(p, q)$.

The properties (1), (2), and (3) appeared (implicitly) in [5].

Proof. By Lemmas 2.7 and 2.14, we can assume that valuation v is equal to the rank function r .

(3). Necessarily $a \preceq p \wedge c$ and $b \preceq q \wedge c$, implying $a \vee b \preceq (p \wedge c) \vee (q \wedge c) \preceq c = a \vee b$. Hence $(p \wedge c) \vee (q \wedge c) = c$. Also $(p \wedge c) \wedge (q \wedge c) = p \wedge q$ (since $p \wedge q \preceq c$). By the modularity equality, we have $r(a) + r(b) = r(c) + r(p \wedge q) = r(p \wedge c) + r(q \wedge c)$, which implies $r[a, p \wedge c] = r[b, q \wedge c] = 0$. Thus $p \wedge c = a$ and $q \wedge c = b$ must hold.

(1). We use the induction on $d(p, q)$. Take a neighbor q' of q in $I(p, q)$. By induction $d(p, q') = r[p \wedge q', p] + r[p \wedge q', q']$, and either (i) q covers q' or (ii) q' covers q . In the first case (i), we must have $p \wedge q \preceq q'$. Suppose not. Then $(p \wedge q) \vee q' = q$, and $(p \wedge q) \wedge q' = p \wedge q'$. The modularity equality yields $r[p \wedge q, q] = r[p \wedge q', q']$, which means that there is a (p, q) -path passing through $p \wedge q$ with the length shorter than $d(p, q')$, contradicting $d(p, q) = d(p, q') + 1$. It follows from $p \wedge q \preceq q'$ that $p \wedge q = p \wedge q'$, and the claim follows. In the second case (ii) where q' covers q , $p \wedge q'$ covers $p \wedge q$; since otherwise

$p \wedge q' = p \wedge q$ which leads to a contradiction $d(p, q') > d(p, q)$. By the modularity equality $r[p \wedge q', q'] = r[p \wedge q, q]$ and the claim follows.

(2). By (1), $p \wedge q \in I(p, q)$. By the modularity equality we have (\supseteq) . We show the reverse inclusion. Take $u \in I(p, q)$. Let $a := u \wedge p$ and $b := u \wedge q$. By (1), $d(p, q) = d(p, u) + d(u, q) = d(p, a) + d(a, u) + d(u, b) + d(b, q)$. Necessarily $u \in I(a, b)$, and $d(a, b) = r[a, u] + r[b, u]$. Since $d(a, b) = r[a \wedge b, a] + r[a \wedge b, b] = r[a, a \vee b] + r[b, a \vee b] = d(a, b) - 2r[a \vee b, u]$, we have $r[a \vee b, u] = 0$, implying $a \vee b = u$. Also $d(p, q) = d(p, a) + d(a, b) + d(b, q)$ must hold. Hence $d(p, q) = d(p, a) + d(a, a \wedge b) + d(a \wedge b, b) + d(b, q) = r[a \wedge b, p] + r[a \wedge b, q] = d(p, q) + 2r[a \wedge b, p \wedge q]$. This implies $a \wedge b = p \wedge q$, and $a \in [p \wedge q, p]$ and $b \in [p \wedge q, q]$.

(4). First we note that $I(p, q)$ is an isometric subspace (with respect to r). Indeed, for $w, w' \in I(p, q)$, by Theorem 2.12, there is a median m of p, w, w' . In particular $m \in I(p, q)$. So $I(w, m) \cup I(m, w') \subseteq I(p, q)$. This means that w and w' are joined by a path in $I(p, q)$ of length $d(w, w')$. By (2) and (3), if w covers w' , then $w \wedge p$ covers $w' \wedge p$ and $w \wedge q = w' \wedge q$, or $w \wedge q$ covers $w' \wedge q$ and $w \wedge p = w' \wedge p$. Thus a shortest path P between u and u' in $I(p, q)$ induces a path P' between $a := u \wedge p$ and $a' := u' \wedge p$ (by map $w \mapsto w \wedge p$) and a path P'' between $b := u \wedge q$ and $b' := u' \wedge q$ (by map $w \mapsto w \wedge q$). The length of P is the sum of lengths of P' and of P'' . This implies that

$$\begin{aligned} d(u, u') &\geq d(a, a') + d(b, b') = r(a) + r(a') - 2r(a \wedge a') + r(b) + r(b') - 2r(b \wedge b') \\ &= r(a \vee b) + r(a' \vee b') - 2r((a \wedge a') \vee (b \wedge b')) \\ &= d(u, u') + 2r[(a \wedge a') \vee (b \wedge b'), u \wedge u'], \end{aligned}$$

where the second equality follows from the modularity equality with $a \wedge b = a' \wedge b' = p \wedge q$ and the third follows from (1). Hence $u \wedge u' = (a \wedge a') \vee (b \wedge b')$, as required. \square

A subset X of \mathcal{L} is called a *subsemilattice* if $p \wedge q \in X$ for any $p, q \in X$, and is called *convex* if X is a convex set in Γ . For an edge-set U , define $\mathcal{L}|U \subseteq \mathcal{L}$ by

$$(2.11) \quad \mathcal{L}|U := \{p \in \mathcal{L} \mid \text{any shortest path from } \mathbf{0} \text{ to } p \text{ belongs to } U\}.$$

Lemma 2.16. (1) *Any convex set in \mathcal{L} is a modular subsemilattice of \mathcal{L} .*

(2) *Suppose that \mathcal{L} is a lattice. Then a subset C is convex if and only if $C = [a, b]$ for some $a, b \in \mathcal{L}$ with $a \leq b$.*

(3) *Suppose that \mathcal{L} is complemented. For an orbit-union U , $\mathcal{L}|U$ is convex, and is a complemented modular subsemilattice of \mathcal{L} . For $p \in \mathcal{L}$, define $p|U \in \mathcal{L}|U$ by*

$$p|U := \text{the gate of } p \text{ at } \mathcal{L}|U.$$

Then $p|U \preceq p$, and any shortest path between p and $p|U$ does not meet U .

Proof. (1) follows from Lemma 2.15. The if part of (2) also follows from Lemma 2.15. To see the only if part of (2), consider $a := \bigwedge_{u \in C} u$ and $b := \bigvee_{u \in C} u$. Then $C \subseteq [a, b]$. From $a, b \in C$, we have $[a, b] = I(a, b) \subseteq C$.

(3). For $p, q \in \mathcal{L}|U$, there is a shortest path from $\mathbf{0}$ to p (or q) passing through $p \wedge q$. This means $[p \wedge q, p], [p \wedge q, q] \subseteq \mathcal{L}|U$. By Lemma 2.15 and Lemma 2.6 (2), it holds that $I(p, q) \subseteq \mathcal{L}|U$. Hence $\mathcal{L}|U$ is convex, and is a modular subsemilattice by (1). Since $[p, q] = I(p, q) \subseteq \mathcal{L}|U$ for $p, q \in \mathcal{L}|U$ with $p \leq q$, every interval of $\mathcal{L}|U$ is complemented.

By $\mathbf{0} \in \mathcal{L}|U$ and the definition of gates, we have $p|U \in I(\mathbf{0}, p) = [\mathbf{0}, p]$. Hence $p|U \preceq p$. Suppose that there is a shortest path from $p|U$ to p having an edge st in U . Suppose that s is covered by t . By the relative complementarity of $[p|U, p]$, there is s' such that $s' \vee s = t$

and $s \wedge s' = p|U$. Then s' covers $p|U$. So $d(p|U, t) = d(s', t) + 1$ and $d(p, p|U) = d(p, s') + 1$ hold. In particular, edges st and $(p|U)s'$ are projective (Lemma 2.5). Thus $(p|U)s' \in U$, and $s' \in \mathcal{L}|U$. By definition of gates, we have $d(p, s') = d(p, p|U) + 1$, contradicting $d(p, p|U) = d(p, s') + 1$. \square

3 Submodular function on modular semilattice

In this section, we develop a theory of submodular functions on modular semilattices. A modular semilattice \mathcal{L} is not necessarily a lattice. Join $p \vee q$ of elements p, q may or may not exist. Interestingly, we can define a certain kind of a join, called a *fractional join*, which is a formal convex combination of elements of a set $\mathcal{E}(p, q) \subseteq \mathcal{L}$ determined by (p, q) :

$$\sum_{u \in \mathcal{E}(p, q)} c(u; p, q)u.$$

If p, q have the join $p \vee q$, then the fractional join is equal to $1(p \vee q)$. In Section 3.1, the set $\mathcal{E}(p, q)$ and the coefficient $c(u; p, q)$ are introduced. Then, in Section 3.2, a function $f : \mathcal{L} \rightarrow \mathbf{R}$ is defined to be a *submodular function* if it satisfies

$$f(p) + f(q) \geq f(p \wedge q) + \sum_{u \in \mathcal{E}(p, q)} c(u; p, q)f(u) \quad (p, q \in \mathcal{L}).$$

The main properties of our submodular functions are:

- The distance function $d = d_{\mathcal{L}}$ on \mathcal{L} is submodular on $\mathcal{L} \times \mathcal{L}$ (Theorem 3.6).
- Submodular functions admit a fractional polymorphism containing semilattice operation \wedge , and hence $\mathbf{VCSP}[\Lambda]$ for submodular language Λ can be solved in polynomial time by the basic LP relaxation (Theorem 3.9).

For readability, less obvious theorems will be proved in Section 3.3.

Although our framework for submodularity was motivated by its application to 0-extension problems, it turned out that several other submodular-type functions, mentioned in the introduction, fall into our framework; see [25, 26] for detail.

3.1 Fractional join

Let \mathcal{L} be a modular semilattice, where its valuation is denoted by v . We begin by sketching the construction of the fractional join of p, q ; see Figure 4. By valuation v and the expression in Lemma 2.15 (3), the metric interval $I(p, q)$ is naturally mapped to the plane \mathbf{R}^2 . Consider the convex hull $\text{Conv } I(p, q)$ of the image of $I(p, q)$. Then the fractional join is a formal sum of elements $u \in I(p, q)$ mapped to maximal extreme points of $\text{Conv } I(p, q)$; the set of such elements is called the (p, q) -*envelope*. The coefficient of u is determined by the normal cone $C(u; p, q)$ at (the image of) u .

(p, q) -envelope. First we introduce the concept of the (p, q) -envelope. Let (p, q) be a pair of elements in \mathcal{L} . Define vector $v(u; p, q)$ in \mathbf{R}_+^2 by

$$(3.1) \quad v(u; p, q) := (v[p \wedge q, u \wedge p], v[p \wedge q, u \wedge q]).$$

Let $\text{Conv } I(p, q)$ denote the convex hull of $\{v(u; p, q) \mid u \in I(p, q)\}$ in \mathbf{R}_+^2 . The polygon $\text{Conv } I(p, q)$ contains $v(p \wedge q; p, q) = (0, 0)$, $v(p; p, q) = (v[p \wedge q, p], 0)$, and $v(q; p, q) = (0, v[p \wedge q, q])$ as extreme points, and contains horizontal segment $[v(p \wedge q; p, q), v(p; p, q)]$

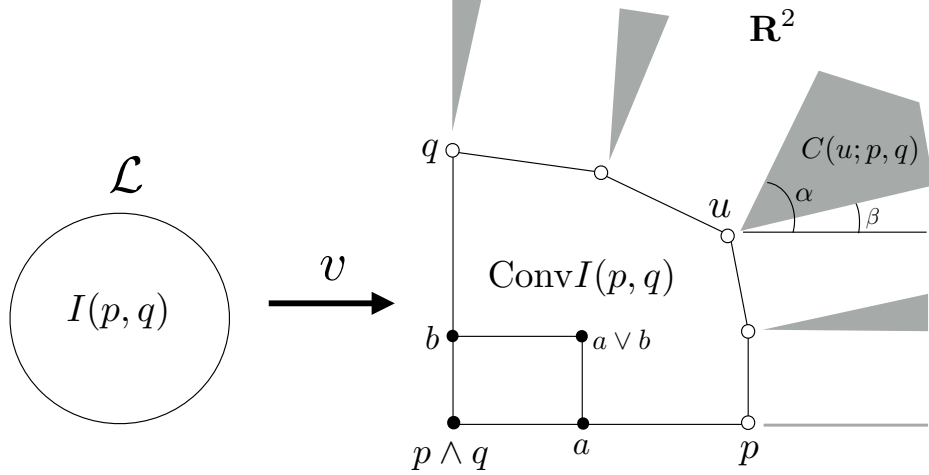


Figure 4: Construction of fractional join

and vertical segment $[v(p \wedge q; p, q), v(q; p, q)]$ as edges. Also $\text{Conv } I(p, q)$ is contained in the rectangle of four vertices

$$(0, 0), (v[p \wedge q, p], 0), (v[p \wedge q, p], 0), (v[p \wedge q, p], v[p \wedge q, q]).$$

The (p, q) -envelope $\mathcal{E}(p, q)$ is the set of elements $u \in I(p, q)$ such that $v(u; p, q)$ is a maximal extreme point of $\text{Conv } I(p, q)$, where a *maximal* extreme point is an extreme point z in $\text{Conv } I(p, q)$ such that for every positive vector ϵ it holds $z + \epsilon \notin \text{Conv } I(p, q)$. Observe that $\mathcal{E}(p, q)$ always contains p and q .

Lemma 3.1. *The map $u \mapsto v(u; p, q)$ is injective on $\mathcal{E}(p, q)$.*

This lemma will be proved in Section 3.3.1. Hence the map $u \mapsto v(u; p, q)$ is a bijection between $\mathcal{E}(p, q)$ and the set of maximal extreme points of $\text{Conv } I(p, q)$.

Valuation of convex cones in \mathbf{R}_+^2 . To define the coefficient, we consider a valuation of convex cones in \mathbf{R}_+^2 . Every closed convex cone $C (\neq \{0\})$ in \mathbf{R}_+^2 is uniquely represented as

$$C = \{(x, y) \in \mathbf{R}_+^2 \mid -x \sin \alpha + y \cos \alpha \leq 0, -x \sin \beta + y \cos \beta \geq 0\}$$

for some $0 \leq \beta \leq \alpha \leq \pi/2$. Define $[C]$ by

$$(3.2) \quad [C] := \frac{\sin \alpha}{\cos \alpha + \sin \alpha} - \frac{\sin \beta}{\cos \beta + \sin \beta}.$$

For convention, we let $[\{0\}] := 0$. The following property is easily verified, where C, C' are closed convex cones in \mathbf{R}_+^2 :

- $$(3.3) \quad \begin{aligned} (1) \quad & [C] \geq 0, \text{ and } [C] > 0 \text{ if and only if } C \text{ is full dimensional.} \\ (2) \quad & [C] + [C'] = [C \cap C'] + [C \cup C'] \text{ for } C \cap C' \neq \emptyset. \\ (3) \quad & [\mathbf{R}_+^2] = 1. \end{aligned}$$

We are now ready to define the fractional join.

Fractional join. For $u \in I(p, q)$, let $C(u; p, q)$ denote the set of nonnegative vectors $w \in \mathbf{R}_+^2$ with $\langle w, v(u; p, q) \rangle = \max_{u' \in I(p, q)} \langle w, v(u'; p, q) \rangle$, where $\langle \cdot, \cdot \rangle$ is the standard inner product. Namely $C(u; p, q)$ is the intersection of \mathbf{R}_+^2 and the normal cone of $\text{Conv } I(p, q)$ at extreme point $v(u; p, q)$. In particular, $C(u; p, q)$ forms a closed convex cone in \mathbf{R}_+^2 . The *fractional join* of (p, q) is the formal convex combination of $u \in \mathcal{E}(p, q)$ with coefficient $[C(u; p, q)]$:

$$\sum_{u \in \mathcal{E}(p, q)} [C(u; p, q)] u.$$

Obviously it holds $[C(u; p, q)] = 0$ for $u \in I(p, q) \setminus \mathcal{E}(p, q)$ since $C(u; p, q) = \{0\}$. So the fractional join is also equal to $\sum_{u \in I(p, q)} [C(u; p, q)] u$. Note that the set $\mathcal{E}(p, q)$ and the coefficient $[C(u; p, q)]$ depend on valuation v . Since the set of cones $C(u; p, q)$ ($u \in \mathcal{E}(p, q)$) forms the intersection of \mathbf{R}_+^2 and the normal fan of $\text{Conv } I(p, q)$, we have

$$(3.4) \quad (1) \quad \mathbf{R}_+^2 = \bigcup_{u \in \mathcal{E}(p, q)} C(u; p, q).$$

(2) For distinct $u, u' \in \mathcal{E}(p, q)$, the intersection $C(u; p, q) \cap C(u'; p, q)$ has no interior point, and hence $[C(u; p, q) \cap C(u'; p, q)] = 0$.

$$(3) \quad \bigcup_{u \in \mathcal{E}(p, q)} [C(u; p, q)] = 1.$$

Therefore the fractional join of p, q is a formal convex combination of elements in $\mathcal{E}(p, q)$. An explicit formula of $[C(u; p, q)]$ is given as follows.

Lemma 3.2. Suppose that $\mathcal{E}(p, q) = \{p = u_0, u_1, u_2, \dots, u_m = q\}$, and $v(u_i; p, q)$ and $v(u_{i+1}; p, q)$ are adjacent extreme points in $\text{Conv } I(p, q)$. Then we have

$$[C(u_i; p, q)] = \delta_i - \delta_{i-1} \quad (i = 0, 1, 2, \dots, m),$$

where δ_i is defined by $\delta_{-1} := 0$, $\delta_m := 1$, and

$$\delta_i := \frac{v[u_i \wedge u_{i+1}, u_i]}{v[u_i \wedge u_{i+1}, u_i] + v[u_i \wedge u_{i+1}, u_{i+1}]} \quad (i = 1, 2, \dots, m-1).$$

This lemma will be proved in Section 3.3.1. We next define the *fractional join operation*. Let $\mathcal{O}(\mathcal{L})$ denote the set of all binary operations $\vartheta : \mathcal{L} \times \mathcal{L} \rightarrow \mathcal{L}$. Regard $\mathcal{O}(\mathcal{L})$ as a poset by the order: $\vartheta \preceq \vartheta'$ if $\vartheta(p, q) \preceq \vartheta'(p, q)$ for all p, q . Then $\mathcal{O}(\mathcal{L})$ is isomorphic to the product $\mathcal{L}^{\mathcal{L} \times \mathcal{L}}$ of \mathcal{L} by the correspondence:

$$(3.5) \quad \mathcal{O}(\mathcal{L}) \ni \vartheta \longleftrightarrow (\vartheta(p, q) : p, q \in \mathcal{L}) \in \mathcal{L}^{\mathcal{L} \times \mathcal{L}}.$$

In particular $\mathcal{O}(\mathcal{L})$ is also a modular semilattice, where the meet \wedge is given by $(\vartheta \wedge \vartheta')(p, q) := \vartheta(p, q) \wedge \vartheta'(p, q)$. The valuation of $\mathcal{O}(\mathcal{L})$ is given according to (2.10), and is also denoted by v . Let L, R be the *projection operations* defined by

$$L(p, q) := p, \quad R(p, q) := q \quad (p, q \in \mathcal{L}).$$

Let $I(\mathcal{L}) := I(L, R)$, and let $\mathcal{E}(\mathcal{L})$ denote the (L, R) -envelope $\mathcal{E}(L, R)$ in $\mathcal{O}(\mathcal{L})$. An operation in $\mathcal{E}(\mathcal{L})$ is called *extremal*. For an operation ϑ in $I(\mathcal{L})$, the cone $C(\vartheta; L, R)$ is denoted simply by $C(\vartheta)$. The *fractional join operation* is the formal sum of extremal operations ϑ with coefficient $[C(\vartheta)]$:

$$\sum_{\vartheta \in \mathcal{E}(\mathcal{L})} [C(\vartheta)] \vartheta.$$

The fractional join operation is nothing but the fractional join of L, R in $\mathcal{O}(\mathcal{L})$, and indeed gives fractional joins in \mathcal{L} .

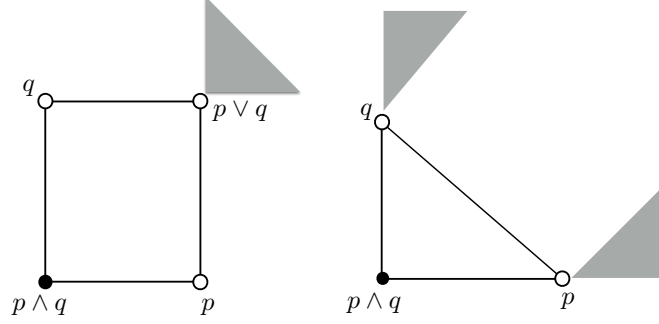


Figure 5: $\text{Conv } I(p, q)$ for a bounded pair (left) and an antipodal pair (right)

Proposition 3.3. $\sum_{u \in \mathcal{E}(p, q)} [C(u; p, q)]u = \sum_{\vartheta \in \mathcal{E}(\mathcal{L})} [C(\vartheta)]\vartheta(p, q) \quad (p, q \in \mathcal{L}).$

The proof is given in Section 3.3.2. Consider the case where \mathcal{L} is the product of modular semilattices \mathcal{L}_i for $i = 1, 2, \dots, n$. For operations ϑ_i on \mathcal{L}_i ($i = 1, 2, \dots, n$), the *componentwise extension* $(\vartheta_1, \vartheta_2, \dots, \vartheta_n)$ is an operation on \mathcal{L} defined by

$$\begin{aligned} (\vartheta_1, \vartheta_2, \dots, \vartheta_n)(p, q) &= (\vartheta_1(p_1, q_1), \vartheta_2(p_2, q_2), \dots, \vartheta_n(p_n, q_n)) \\ &\quad (p, q \in \mathcal{L} = \mathcal{L}_1 \times \mathcal{L}_2 \times \dots \times \mathcal{L}_n). \end{aligned}$$

Proposition 3.4.

$$\sum_{\vartheta \in \mathcal{E}(\mathcal{L})} [C(\vartheta)]\vartheta = \sum_{\vartheta_1, \vartheta_2, \dots, \vartheta_n} [C(\vartheta_1) \cap C(\vartheta_2) \cap \dots \cap C(\vartheta_n)](\vartheta_1, \vartheta_2, \dots, \vartheta_n),$$

where ϑ_i is taken over all extremal operations in \mathcal{L}_i ($i = 1, 2, \dots, n$). Moreover, if $\mathcal{L}_i = \mathcal{L}_j$ and $\vartheta_i \neq \vartheta_j$ for some i, j , then $[C(\vartheta_1) \cap C(\vartheta_2) \cap \dots \cap C(\vartheta_n)] = 0$.

The proof is given in Section 3.3.3. In particular, any extremal operation ϑ in \mathcal{L} is the componentwise extension of extremal operations θ_i in \mathcal{L}_i for $i = 1, 2, \dots, n$.

3.2 Submodular function

Let \mathcal{L} be a modular semilattice with valuation v . A function $f : \mathcal{L} \rightarrow \overline{\mathbf{R}}$ is called *submodular* on \mathcal{L} (with respect to v) if it satisfies

$$(3.6) \quad f(p) + f(q) \geq f(p \wedge q) + \sum_{u \in \mathcal{E}(p, q)} [C(u; p, q)]f(u) \quad (p, q \in \mathcal{L}).$$

By Proposition 3.3, a submodular function may also be characterized as a function f satisfying

$$(3.7) \quad f(p) + f(q) \geq f(p \wedge q) + \sum_{\vartheta \in \mathcal{E}(\mathcal{L})} [C(\vartheta)]f(\vartheta(p, q)) \quad (p, q \in \mathcal{L}).$$

In the case where a pair (p, q) is bounded, the join (p, q) exists, $\text{Conv } I(p, q)$ is a square of vertices $(0, 0)$, $(0, v[p \wedge q, q])$, $(v[p \wedge q, p], 0)$, $(v[p \wedge q, p], v[p \wedge q, q])$, $\mathcal{E}(p, q) = \{p, p \vee q, q\}$, and the fractional join is equal to $0p + 1(p \vee q) + 0q = p \vee q$. See Figure 5. Hence the corresponding inequality in (3.6) is equal to the usual submodularity inequality:

$$(3.8) \quad f(p) + f(q) \geq f(p \wedge q) + f(p \vee q).$$

Another extremal case is the case of $\mathcal{E}(p, q) = \{p, q\}$, which implies that $\text{Conv } I(p, q)$ is the triangle with vertices $(0, 0)$, $(v[p \wedge q, p], 0)$, $(0, v[p \wedge q, q])$. Such a pair (p, q) is called *antipodal*. By definition, a pair (p, q) is antipodal if and only if every bounded pair (a, b) with $p \succeq a \succeq p \wedge q \preceq b \preceq q$ satisfies

$$(3.9) \quad v[a, p]v[b, q] \geq v[p \wedge q, a]v[p \wedge q, b].$$

This condition rephrases that the point $v(a \vee b; p, q)$ is lower than the line through the points $v(p; p, q)$ and $v(q; p, q)$. In this case, the fractional join of (p, q) is equal to

$$\frac{v[p \wedge q, p]}{v[p \wedge q, p] + v[p \wedge q, q]}p + \frac{v[p \wedge q, q]}{v[p \wedge q, p] + v[p \wedge q, q]}q.$$

The inequality in (3.6) corresponds to

$$(3.10) \quad v[p \wedge q, q]f(p) + v[p \wedge q, p]f(q) \geq (v[p \wedge q, p] + v[p \wedge q, q])f(p \wedge q).$$

We call this inequality the \wedge -convexity inequality. Then inequalities (3.8) and (3.10) suffice to characterize the submodularity:

Theorem 3.5. *$f : \mathcal{L} \rightarrow \overline{\mathbf{R}}$ is submodular if and only if it satisfies*

- (1) $\mathcal{E}(p, q) \subseteq \text{dom } f$ for $p, q \in \text{dom } f$,
- (2) the submodularity inequality for every bounded pair (p, q) , and
- (3) the \wedge -convexity inequality for every antipodal pair (p, q) .

The proof is given in Section 3.3.1. One of the main properties of our submodularity is the following:

Theorem 3.6. *Let \mathcal{L} be a modular semilattice. The distance function $d_{\mathcal{L}}$ on \mathcal{L} is submodular on $\mathcal{L} \times \mathcal{L}$.*

We will prove this theorem and a more general version (Theorem 4.8) in Section 4.3.3. We note some basic properties of our submodular functions concerning addition and restriction.

Lemma 3.7. *Let $\mathcal{L}, \mathcal{L}', \mathcal{M}$ and \mathcal{M}' be modular semilattices, and let f and f' be submodular functions on \mathcal{L} .*

- (1) For $b \in \mathbf{R}$ and $c, c' \in \mathbf{R}_+$, $b + cf + c'f'$ is a submodular function on \mathcal{L} .
- (2) A function \tilde{f} defined by

$$\tilde{f}(p, p') := f(p) \quad ((p, p') \in \mathcal{L} \times \mathcal{L}')$$

is a submodular function on $\mathcal{L} \times \mathcal{L}'$.

- (3) Suppose $\mathcal{L} = \mathcal{M} \times \mathcal{M}'$. For any $p' \in \mathcal{M}'$, a function $f_{p'}$ defined by

$$f_{p'}(p) := f(p, p') \quad (p \in \mathcal{M})$$

is a submodular function on \mathcal{M} .

- (4) For a convex set \mathcal{N} of \mathcal{L} , the restriction of f to \mathcal{N} is a submodular function on \mathcal{N} (regarded as a modular semilattice).

Proof. (1) follows from the facts that the submodularity is closed under nonnegative sum, and that any constant function is submodular (by (3.4) (3)).

(2) follows from (3.7), Proposition 3.4, and

$$\begin{aligned}\tilde{f}(p, p') + \tilde{f}(q, q') &= f(p) + f(q) \geq f(p \wedge q) + \sum_{\vartheta \in \mathcal{E}(\mathcal{L})} [C(\vartheta)]f(\vartheta(p, q)) \\ &= \tilde{f}(p \wedge q, p' \wedge q') + \sum_{\vartheta \in \mathcal{E}(\mathcal{L})} \sum_{\vartheta' \in \mathcal{E}(\mathcal{L}')} [C(\vartheta) \cap C(\vartheta')] \tilde{f}(\vartheta(p, q), \vartheta'(p', q')), \end{aligned}$$

where we use $[C(\vartheta)] = \sum_{\vartheta' \in \mathcal{E}(\mathcal{L}')} [C(\vartheta) \cap C(\vartheta')]$ (obtained from (3.4)).

(3). Notice that the (p, p) -envelope is $\{p\}$, and any extremal operation ϑ is idempotent, i.e., $\vartheta(p, p) = p$. Thus, by Proposition 3.4, we have

$$\begin{aligned}f_{p'}(p) + f_{p'}(q) &= f(p, p') + f(q, p') \\ &\geq f(p \wedge q, p' \wedge p') + \sum_{\vartheta \in \mathcal{E}(\mathcal{M})} \sum_{\vartheta' \in \mathcal{E}(\mathcal{M}')} [C(\vartheta) \cap C(\vartheta')] f(\vartheta(p, q), \vartheta'(p', p')) \\ &= f(p \wedge q, p') + \sum_{\vartheta \in \mathcal{E}(\mathcal{M})} \sum_{\vartheta' \in \mathcal{E}(\mathcal{M}')} [C(\vartheta) \cap C(\vartheta')] f(\vartheta(p, q), p') \\ &= f_{p'}(p \wedge q) + \sum_{\vartheta \in \mathcal{E}(\mathcal{M})} [C(\vartheta)] f_{p'}(\vartheta(p, q)). \end{aligned}$$

(4) follows from the fact that for $p, q \in \mathcal{N}$ the metric interval $I(p, q)$ is the same on \mathcal{L} and on \mathcal{N} . \square

We finally give a useful criterion of the submodularity. A bounded pair (p, q) in \mathcal{L} is said to be *2-bounded* if $p \vee q$ covers p and q (in which case both p and q cover $p \wedge q$).

Proposition 3.8. *Suppose that \mathcal{L} is the product of two modular semilattices $\mathcal{L}_1 \times \mathcal{L}_2$. $f : \mathcal{L} \rightarrow \mathbf{R}$ is submodular if and only if it satisfies*

- (1) *the submodularity inequality for every 2-bounded pair p, q , and*
- (2) *the \wedge -convexity inequality for every pair $(p, q) = ((p_1, p_2), (q_1, q_2))$ such that $p_1 = q_1$ and (p_2, q_2) is antipodal in \mathcal{L}_2 or $p_2 = q_2$ and (p_1, q_1) is antipodal in \mathcal{L}_1 .*

The proof is given in Section 3.3.4. Notice that this criterion does not work when f has infinite values.

Minimizing a sum of submodular functions with bounded arity. Here we consider the problem of minimizing submodular function f on the product of modular semilattices $\mathcal{L}_1, \mathcal{L}_2, \dots, \mathcal{L}_n$, where the input of the problem is $\mathcal{L}_1, \mathcal{L}_2, \dots, \mathcal{L}_n$ and an evaluating oracle of f . In the case where each \mathcal{L}_i is a lattice of rank 1, this problem is the submodular set function minimization in the ordinary sense, and can be solved in polynomial time [20, 30, 54]. However, we do not know whether this problem in general is polynomial time solvable or not. One notable result in this direction, due to Kuivinen [43], is that if each \mathcal{L}_i is a complemented modular lattice of rank 2 (a *diamond lattice*), then this problem has a *good characterization*.

So we restrict our investigation to the problem of minimizing a sum of submodular functions with bounded arity, i.e., valued CSP for submodular functions. See Section 2.1 for notions in valued CSP. Let $\mathcal{L} := \mathcal{L}_1 \times \mathcal{L}_2 \times \dots \times \mathcal{L}_n$. A constraint f on \mathcal{L} is called *submodular* if f is a submodular function on \mathcal{L}_{I_f} . The *submodular language* $\mathcal{S}_{\mathcal{L}}$ is the set of all submodular constraints on \mathcal{L} .

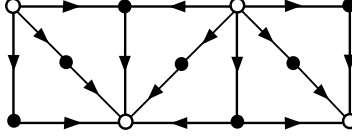


Figure 6: A nonsemilattice orientable modular graph

Theorem 3.9. *Let \mathcal{L} be the product of modular semilattices $\mathcal{L}_1, \mathcal{L}_2, \dots, \mathcal{L}_n$. Then $\mathbf{VCSP}[\mathcal{S}_{\mathcal{L}}]$ can be solved in polynomial time.*

Indeed, the submodular language $\mathcal{S}_{\mathcal{L}}$ satisfies the Thapper-Živný criterion (Theorem 2.1). Define a fractional operation ω on \mathcal{L} by

$$(3.11) \quad \omega = \frac{1}{2} \wedge + \frac{1}{2} \sum_{\vartheta \in \mathcal{E}(\mathcal{L})} [C(\vartheta)] \vartheta.$$

By Proposition 3.4, an extremal operation is the componentwise extension of operations in \mathcal{L}_i , and is separable. Also, by (3.4) (3), the total sum of coefficients $[C(\vartheta)]$ is equal to 1. Therefore ω is a fractional polymorphism for the submodular language $\mathcal{S}_{\mathcal{L}}$. Obviously \wedge is a semilattice operation. Hence Theorem 3.9 follows from Theorem 2.1.

Remark 3.10. Suppose that Γ is the covering graph of a modular semilattice \mathcal{L} . By Theorem 3.6 and Lemma 3.7, g_j and f_{ij} (defined in (2.1)) are submodular constraints on \mathcal{L}^n . Then $\mathbf{0-Ext}[\Gamma]$ is a subclass of $\mathbf{VCSP}[\mathcal{S}_{\mathcal{L}^n}]$. Therefore, by Theorem 3.9, $\mathbf{0-Ext}[\Gamma]$ can be solved by the basic LP relaxation. This observation, however, does not get us to the main result (Theorem 1.6) since there is an orientable modular graph that cannot be represented as the covering graph of a modular semilattice. See the graph of Figure 6, where there are exactly two admissible orientations: one is the reverse of the other, and both orientations have two sinks and two sources. Nevertheless, the basic LP is expected to solve $\mathbf{0-Ext}[\Gamma]$ for an arbitrary orientable modular graph Γ ; see Section 6.

3.3 Proofs

3.3.1 Proof of Lemmas 3.1 and 3.2 and Theorem 3.5

Let (p, q) be a pair of elements in modular semilattice \mathcal{L} . First we prove a general version of Lemma 3.1.

Lemma 3.11. *For $s, s' \in \mathcal{E}(p, q)$, if $v(s \wedge p) \geq v(s' \wedge p)$ and $v(s \wedge q) \leq v(s' \wedge q)$, then $s \wedge p \succeq s' \wedge p$ and $s \wedge q \preceq s' \wedge q$. In particular, $v(s; p, q) = v(s'; p, q)$ implies $s = s'$.*

Proof. We can assume that $v(p \wedge q) = 0$ by adding a constant to v (for notational simplicity). It suffices to consider the case where $v(s; p, q) = v(s'; p, q)$ or $v(s; p, q)$ and $v(s'; p, q)$ are adjacent extreme points in $\text{Conv } I(p, q)$. Let $(a, b) := (s \wedge p, s \wedge q)$ and $(a', b') := (s' \wedge p, s' \wedge q)$. All pairs (a, a') , $(a, b \wedge b')$, and $(a', b \wedge b')$ from among the triple $(a, a', b \wedge b')$ are bounded. By definition of modular semilattices, their join $\eta := (a \vee a') \vee (b \wedge b')$ exists and belongs to $I(p, q)$ (see Lemma 2.15). Similarly the join $\xi := (a \wedge a') \vee (b \vee b')$ of triple $(a \wedge a', b, b')$ exists and belongs to $I(p, q)$. Therefore $v(\eta; p, q) = (v(a \vee a'), v(b \wedge b'))$ and $v(\xi; p, q) = (v(a \wedge a'), v(b \vee b'))$. By modularity equality (2.7) for v , we have

$$v(s; p, q) + v(s'; p, q) = (v(a) + v(a'), v(b) + v(b')) = v(\eta; p, q) + v(\xi; p, q).$$

Then both $v(\eta; p, q)$ and $v(\xi; p, q)$ must belong to $[v(u; p, q), v(u'; p, q)]$ since it is an edge (or an extreme point) of $\text{Conv } I(p, q)$. Necessarily $v(\eta; p, q) = v(s; p, q)$ and $v(\xi; p, q) = v(s'; p, q)$. This means that $a \vee a' = a$ and $b \vee b' = b'$. Hence the claim follows. \square

Lemma 3.12. *For $s, t \in \mathcal{E}(p, q)$ with $t \wedge p \preceq s \wedge p$ (and $t \wedge q \succeq s \wedge q$), the following hold:*

- (1) $d(p, q) = d(p, s) + d(s, t) + d(t, q)$; in particular $I(s, t) \subseteq I(p, q)$.
- (2) $v(u; p, q) = v(u; s, t) + v(s \wedge t; p, q)$ for $u \in I(s, t) \subseteq I(p, q)$.
- (3) If $v(s; p, q)$ and $v(t; p, q)$ are adjacent extreme points, then (s, t) is antipodal.

Proof. (1). By Lemma 2.15 (4), we have $(s \wedge t) \vee (q \wedge t) = (s \wedge t \wedge p) \vee (s \wedge t \wedge q) \vee (q \wedge t) = (t \wedge p) \vee (t \wedge q) = t$. This means $t \in I(s, q)$. Hence $d(p, q) = d(p, s) + d(s, q) = d(p, s) + d(s, t) + d(t, q)$.

(2). By $u \wedge s = (s \wedge t) \vee (u \wedge p)$ (Lemma 2.15 (4) for $I(p, t)$), we have $v[t \wedge p, p \wedge u] = v[s \wedge t, u \wedge s]$, and

$$v[p \wedge q, p \wedge u] = v[p \wedge q, t \wedge p] + v[t \wedge p, p \wedge u] = v[p \wedge q, s \wedge t \wedge p] + v[s \wedge t, u \wedge s].$$

Similarly (for $I(s, q)$), we have $v[p \wedge q, q \wedge u] = v[p \wedge q, s \wedge t \wedge q] + v[s \wedge t, u \wedge t]$.

(3). If (s, t) is not antipodal, then there is $u \in I(s, t)$ such that $v(u; s, t)$ goes beyond the line segment between $v(s; s, t)$ and $v(t; s, t)$. Then, by (2), $v(u; p, q)$ is in the outside of $\text{Conv } I(p, q)$. This is a contradiction. \square

Suppose that $\mathcal{E}(p, q) = \{p = u_0, u_1, \dots, u_m = q\}$, and $v(u_i; p, q)$ and $v(u_{i+1}; p, q)$ are adjacent extreme points. Let θ_i be the angle of the line normal to the line segment connecting $v(u_{i-1}; p, q)$ and $v(u_i; p, q)$. By Lemma 3.12 (2), θ_i is equal to the angle of the line normal to the line segment connecting $v(u_{i-1}; u_i, u_{i-1})$ and $v(u_i; u_i, u_{i-1})$. Therefore

$$(3.12) \quad \frac{\sin \theta_i}{\sin \theta_i + \cos \theta_i} = \frac{v[u_i \wedge u_{i-1}, u_{i-1}]}{v[u_i \wedge u_{i-1}, u_{i-1}] + v[u_i \wedge u_{i-1}, u_i]} = \delta_{i-1}.$$

Therefore we obtain the formula of Lemma 3.2.

Next we prove Theorem 3.5. It suffices to prove the if part. Let (p, q) be a pair of (incomparable) elements in \mathcal{L} . Suppose that $\mathcal{E}(p, q) = \{p = u_0, u_1, \dots, u_m = q\}$ is given as above. Let $p_i := u_i \wedge p$ and $q_i := u_i \wedge q$ for $i = 0, 1, 2, \dots, m$. By Lemma 3.11, it holds $p_i \succeq p_j$ and $q_i \preceq q_j$ for $i \leq j$. Let $b_i := u_{i-1} \wedge u_i$ for $i = 1, 2, \dots, m$; see Figure 7. Then we have

$$(3.13) \quad b_i \wedge q_i = u_{i-1} \wedge u_i \wedge u_i \wedge q = q_{i-1}.$$

Also, by Lemma 2.15 (3) and (4), we have

$$(3.14) \quad \begin{aligned} b_i \vee q_i &= (u_{i-1} \wedge u_i) \vee q_i = (u_{i-1} \wedge u_i \wedge p) \vee (u_{i-1} \wedge u_i \wedge q) \vee q_i \\ &= (p_{i-1} \wedge p_i) \vee (q_{i-1} \wedge q_i) \vee q_i = p_i \vee q_i = u_i. \end{aligned}$$

Let f be a function on \mathcal{L} satisfying the conditions (1), (2), and (3) in Theorem 3.5. We show that f satisfies the inequality (3.6) for p, q . We may assume that $p, q \in \text{dom } f$. By condition (1), $\mathcal{E}(p, q) \subseteq \text{dom } f$. Namely $u_i \in \text{dom } f$. By (3), we have $b_i = u_i \wedge u_{i-1} \in \text{dom } f$. By (2), we have $b_i \wedge b_{i-1} = u_i \wedge u_{i-1} \wedge u_{i-2} = u_i \wedge u_{i-2} \in \text{dom } f$. Consequently all p_i, q_i, b_i belong to $\text{dom } f$. By (3.13), (3.14), and the condition (2) (submodularity), we have

$$(3.15) \quad f(b_i) + f(q_i) \geq f(q_{i-1}) + f(u_i) \quad (i = 1, 2, \dots, m).$$

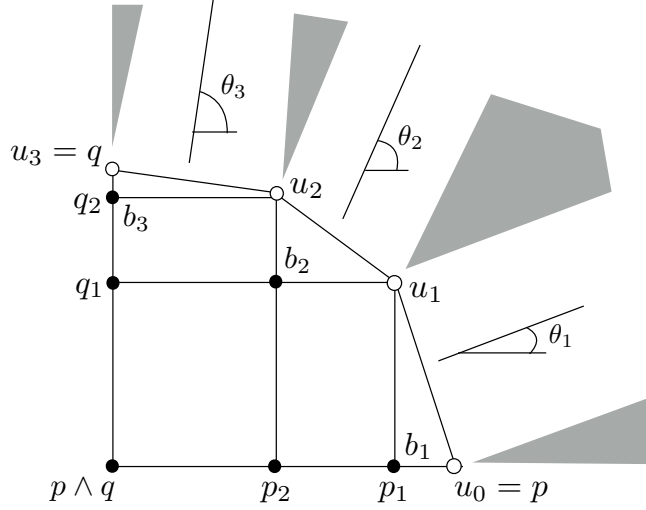


Figure 7: $\text{Conv } I(p, q)$ and $\mathcal{E}(p, q)$

By adding (3.15) for $i = 1, 2, \dots, m$ and $f(p) = f(u_0)$ (recall $(p, q) = (u_0, q_m)$), we obtain

$$(3.16) \quad f(p) + f(b_1) + f(b_2) + \dots + f(b_m) + f(q) \geq f(p \wedge q) + f(u_0) + f(u_1) + \dots + f(u_m).$$

By Lemma 3.12 (3), pair (u_i, u_{i+1}) is antipodal. By condition (3), f satisfies the \wedge -convexity inequality (3.10) for (u_i, u_{i+1}) , which is rewritten as

$$(3.17) \quad f(u_i) \geq f(b_{i+1}) + \delta_i f(u_i) - \delta_i f(u_{i+1}).$$

Substituting (3.17) to (3.16) for $i = 0, 1, 2, \dots, m-1$, we obtain

$$f(p) + f(q) \geq f(p \wedge q) + \sum_{i=0}^m (\delta_i - \delta_{i-1}) f(u_i) = f(p \wedge q) + \sum_{i=0}^m [C(u_i; p, q)] f(u_i).$$

3.3.2 Proof of Proposition 3.3

We start with preliminary arguments. Suppose that \mathcal{L} is the product $\mathcal{L}_1 \times \mathcal{L}_2$ of modular semilattices \mathcal{L}_1 and \mathcal{L}_2 . The valuations of \mathcal{L}_1 and \mathcal{L}_2 are given as (2.8). They are also denoted by v . Let $(p, q) = ((p_1, p_2), (q_1, q_2))$ be a pair of elements in \mathcal{L} . By (2.9), we have

$$(3.18) \quad v(u; p, q) = v(u_1; p_1, q_1) + v(u_2; p_2, q_2) \quad (u \in I(p, q)).$$

By this equation together with $I(p, q) = I(p_1, q_1) \times I(p_2, q_2)$, we have

$$(3.19) \quad \text{Conv } I(p, q) = \text{Conv } I(p_1, q_1) + \text{Conv } I(p_2, q_2),$$

where the sum means the Minkowski sum.

In general, if a polytope P is the Minkowski sum of two polytopes Q and Q' , then every extreme point of P is uniquely represented as the sum of extreme points of Q and of Q' . By this fact and the injectivity of $v(\cdot; p, q)$ (Lemma 3.1) on $\mathcal{E}(p, q)$, we obtain:

$$(3.20) \quad \text{For } u = (u_1, u_2) \in \mathcal{E}(p, q), \text{ there uniquely exist maximal extreme points } x_i \in \text{Conv } I(p_i, q_i) \text{ for } i = 1, 2 \text{ such that } x_i = v(u_i; p_i, q_i) \text{ for } i = 1, 2, \text{ and } v(u; p, q) = x_1 + x_2. \text{ In particular, } u_i \text{ belongs to } \mathcal{E}(p_i, q_i) \text{ for } i = 1, 2.$$

Moreover, $v(u; p, q)$ maximizes $\langle c, x \rangle$ over $x \in \text{Conv } I(p, q)$ if and only if $v(u_i; p_i, q_i)$ maximizes $\langle c, x \rangle$ over $x \in \text{Conv } I(p_i, q_i)$ for $i = 1, 2$. Therefore we have

$$(3.21) \quad C(u; p, q) = C(u_1; p_1, q_1) \cap C(u_2; p_2, q_2).$$

We are ready to prove Proposition 3.3. Regard $\mathcal{O}(\mathcal{L})$ as a modular semilattice $\mathcal{L}^{\mathcal{L} \times \mathcal{L}}$. Then $I(\mathcal{L})$ is the product of $I(L(p, q), R(p, q)) = I(p, q)$ over all $(p, q) \in \mathcal{L} \times \mathcal{L}$. By (3.19) we have

$$\text{Conv } I(\mathcal{L}) = \sum_{(p, q) \in \mathcal{L} \times \mathcal{L}} \text{Conv } I(p, q).$$

By (3.20), for every extremal operation ϑ , each $\vartheta(p, q)$ belongs to $\mathcal{E}(p, q)$. Also, by (3.21), we have

$$(3.22) \quad C(\vartheta) = \bigcap_{(p, q) \in \mathcal{L} \times \mathcal{L}} C(\vartheta(p, q); p, q) \quad (\vartheta \in \mathcal{E}(\mathcal{L})).$$

From (3.4), $C(\vartheta) \subseteq C(u; p, q)$ if and only if $\vartheta(p, q) = u$, and

$$(3.23) \quad C(u; p, q) = \bigcup_{\vartheta \in \mathcal{E}(\mathcal{L}): \vartheta(p, q) = u} C(\vartheta) \quad (p, q \in \mathcal{L}, u \in \mathcal{E}(p, q)),$$

where any two of the cones in the union have no common interior points. Therefore

$$\sum_{\vartheta \in \mathcal{E}(\mathcal{L})} [C(\vartheta)] \vartheta(p, q) = \sum_{u \in \mathcal{E}(p, q)} \left(\sum_{\vartheta \in \mathcal{E}(\mathcal{L}): \vartheta(p, q) = u} [C(\vartheta)] \right) u = \sum_{u \in \mathcal{E}(p, q)} [C(u; p, q)] u.$$

This proves Proposition 3.3.

3.3.3 Proof of Proposition 3.4

It suffices to consider the case where $\mathcal{L} = \mathcal{L}_1 \times \mathcal{L}_2$. If $\vartheta \in I(\mathcal{L})$ is the componentwise extension of $\vartheta_i \in I(\mathcal{L}_i)$ for $i = 1, 2$, then

$$\begin{aligned} (3.24) \quad v(\vartheta; L, R) &= \sum_{(p, q) \in \mathcal{L} \times \mathcal{L}} v(\vartheta(p, q); p, q) \\ &= \sum_{(p, q) \in \mathcal{L} \times \mathcal{L}} v(\vartheta_1(p_1, q_1); p_1, q_1) + v(\vartheta_2(p_2, q_2); p_2, q_2) \\ &= |\mathcal{L}_2|^2 v(\vartheta_1; L, R) + |\mathcal{L}_1|^2 v(\vartheta_2; L, R). \end{aligned}$$

Therefore it holds

$$(3.25) \quad \text{Conv } I(\mathcal{L}) \supseteq |\mathcal{L}_2|^2 \text{Conv } I(\mathcal{L}_1) + |\mathcal{L}_1|^2 \text{Conv } I(\mathcal{L}_2).$$

We are going to show that every extremal operation in \mathcal{L} is the componentwise extension of operations in \mathcal{L}_i for $i = 1, 2$, and the equality holds in (3.25). Take an extremal operation ϑ on \mathcal{L} . There are $\vartheta_i : \mathcal{L} \times \mathcal{L} \rightarrow \mathcal{L}_i$ for $i = 1, 2$ such that $\vartheta(p, q) = (\vartheta_1(p, q), \vartheta_2(p, q))$ for $(p, q) \in \mathcal{L} \times \mathcal{L}$. By Proposition 3.3, $\vartheta(p, q)$ belongs to $\mathcal{E}(p, q)$. By (3.20), $\vartheta_1(p, q)$ and $\vartheta_2(p, q)$ belong to $\mathcal{E}(p_1, q_1)$ and $\mathcal{E}(p_2, q_2)$, respectively. For $(p, q), (p', q') \in \mathcal{L} \times \mathcal{L}$ with $(p_1, q_1) = (p'_1, q'_1)$, suppose (indirectly) that $\vartheta_1(p, q) \neq \vartheta_1(p', q')$. Let ϑ' be the operation

in $I(\mathcal{L})$ obtained from ϑ by replacing $\vartheta_1(p', q')$ with $\vartheta_1(p, q)$, and let ϑ'' be the operation in $I(\mathcal{L})$ obtained from ϑ by replacing $\vartheta_1(p, q)$ with $\vartheta_1(p', q')$. Then we have

$$\begin{aligned} v(\vartheta(p, q); p, q) + v(\vartheta(p', q'); p', q') &= \sum_{i=1,2} v(\vartheta_i(p, q); p_i, q_i) + v(\vartheta_i(p', q'); p'_i, q'_i) \\ &= \frac{1}{2} \{ (v(\vartheta'(p, q); p, q) + v(\vartheta'(p', q'); p', q')) + (v(\vartheta''(p, q); p, q) + v(\vartheta''(p', q'); p', q')) \}. \end{aligned}$$

By Lemma 3.1, $v(\vartheta_1(p, q); p_1, q_1)$ and $v(\vartheta_1(p', q'); p_1, q_1)$ are distinct, and consequently $v(\vartheta; L, R)$ is the midpoint of segment between distinct points $v(\vartheta'; L, R)$ and $v(\vartheta''; L, R)$, contradicting the fact that ϑ is extremal. Therefore $\vartheta_1(p, q) = \vartheta_1(p', q')$ must hold. This means that $\vartheta_1(p, q)$ does not depend on the second component of each of p, q . So we can regard $\vartheta_1 \in I(\mathcal{L}_1)$. Similarly $\vartheta_2 \in I(\mathcal{L}_2)$, and ϑ is equal to the componentwise extension of ϑ_1 and ϑ_2 . Hence the equality holds in (3.25), both ϑ_1 and ϑ_2 must be extremal, and $[C(\vartheta)] = [C(\vartheta_1) \cap C(\vartheta_2)]$. Thus we have

$$\sum_{\vartheta \in \mathcal{E}(\mathcal{L})} [C(\vartheta)]\vartheta = \sum_{\vartheta_1 \in \mathcal{E}(\mathcal{L}_1), \vartheta_2 \in \mathcal{E}(\mathcal{L}_2)} [C(\vartheta_1) \cap C(\vartheta_2)](\vartheta_1, \vartheta_2).$$

Suppose that $\mathcal{L}_1 = \mathcal{L}_2$ holds. Suppose that ϑ_1 and ϑ_2 are different. We see from (3.24) that $v((\vartheta_1, \vartheta_2); L, R)$ is the midpoint of the segment between distinct points $v((\vartheta_1, \vartheta_1); L, R)$ and $v((\vartheta_2, \vartheta_2); L, R)$. Hence $(\vartheta_1, \vartheta_2)$ is never extremal. The proof of Proposition 3.4 is now complete.

3.3.4 Proof of Proposition 3.8

We use the characterization of Theorem 3.5. So the only if part is obvious. We prove the if part. We first show that the submodularity inequality for an arbitrary bounded pair is implied by submodularity inequalities for 2-bounded pairs. For a bounded pair (p, q) , take maximal chains $(p \wedge q = p_0, p_1, \dots, p_k = p)$ and $(p \wedge q = q_0, q_1, \dots, q_l = q)$. Let $a_{i,j} := p_i \vee q_j$. Then $f(p) + f(q) - f(p \wedge q) - f(p \vee q) = \sum_{i,j} (f(a_{i+1,j}) + f(a_{i,j+1}) - f(a_{i+1,j+1}) - f(a_{i,j})) \geq 0$. Here we use the fact seen from modularity that $(a_{i+1,j}, a_{i,j+1})$ is a 2-bounded pair with $a_{i+1,j+1} = a_{i+1,j} \vee a_{i,j+1}$ and $a_{i,j} = a_{i+1,j} \wedge a_{i,j+1}$.

Next we show the \wedge -convexity inequality. Take an (incomparable) antipodal pair $(p, q) = ((p_1, p_2), (q_1, q_2))$ in \mathcal{L} . Then $\mathcal{E}(p, q) = \{p, q\}$. Then $\text{Conv } I(p, q)$ is a triangle. By (3.19), it holds $\text{Conv } I(p, q) = \text{Conv } I(p_1, q_1) + \text{Conv } I(p_2, q_2)$. Therefore both $\text{Conv } I(p_1, q_1)$ and $\text{Conv } I(p_2, q_2)$ are triangles congruent to a dilation of $\text{Conv } I(p, q)$. Hence (p_i, q_i) is antipodal in \mathcal{L}_i , and $C(p_i; p_i, q_i) = C(p; p, q)$ and $C(q_i; p_i, q_i) = C(q; p, q)$ for $i = 1, 2$. In particular, both $((q_1, q_2), (q_1, p_2))$ and $((q_1, p_2), (p_1, p_2))$ are antipodal pairs in $\mathcal{L} = \mathcal{L}_1 \times \mathcal{L}_2$. Letting $C_p := C(p; p, q) = C(p_i; p_i, q_i)$ and $C_q := C(q; p, q) = C(q_i; p_i, q_i)$, we have

$$\begin{aligned} (1 - [C_q])f(q_1, q_2) + (1 - [C_p])f(q_1, p_2) &\geq f(q_1, p_2 \wedge q_2), \\ (1 - [C_q])f(q_1, p_2) + (1 - [C_p])f(p_1, p_2) &\geq f(p_1 \wedge q_1, p_2). \end{aligned}$$

Also, by submodularity inequality (shown above), we have

$$f(q_1, p_2 \wedge q_2) + f(p_1 \wedge q_1, p_2) \geq f(p_1 \wedge q_1, p_2 \wedge q_2) + f(q_1, p_2).$$

From the three inequalities, we obtain

$$\begin{aligned} (1 - [C_q])f(q_1, q_2) + (2 - [C_p] - [C_q])f(q_1, p_2) + (1 - [C_p])f(p_1, p_2) \\ \geq f(p_1 \wedge q_1, p_2 \wedge q_2) + f(q_1, p_2). \end{aligned}$$

By using $[C_p] + [C_q] = 1$, we obtain the \wedge -convexity inequality for (p, q) .

4 L-convex function on modular complex

A *modular complex* $\mathbf{\Gamma}$ is a triple (Γ, o, h) of an orientable modular graph Γ , its admissible orientation o , and its positive orbit-invariant function h . The goal of this section is to introduce a class of discrete convex functions, called *L-convex functions*, on $\mathbf{\Gamma}$, and show that L-convex functions have several nice properties for optimization, analogous to L^\natural -convex functions in discrete convex analysis.

The main properties of our L-convex functions are:

- The distance function $d_{\Gamma, h}$ is an L-convex function on $\mathbf{\Gamma} \times \mathbf{\Gamma}$ (Theorem 4.8).
- In the minimization of an L-convex function, checking optimality and finding a descent direction can be done by submodular function minimization on modular semilattices (Theorem 4.11).

In Section 4.1, we explore several structural properties of modular complexes. In particular, a modular complex can be regarded as a structure obtained by gluing modular semilattices (Theorem 4.2), and admits a subdivision operation (Theorem 4.3). This operation produces a fine modular complex $\mathbf{\Gamma}^*$ into which the original modular complex $\mathbf{\Gamma}$ is embedded, and also enables us to define the *neighborhood semilattice* \mathcal{L}_p^* around each vertex p , which is also a modular semilattice. Based on this investigation as well as the idea mentioned in the introduction, in Section 4.2, we introduce L-convex functions on $\mathbf{\Gamma}$, and present their properties. Again less obvious theorems will be proved in Section 4.3. A further geometric study on orientable modular graphs is given in [10].

4.1 Modular complex

Let $\mathbf{\Gamma} = (\Gamma, o, h)$ be a modular complex, where a modular complex is denoted by the bold style $\mathbf{\Gamma}$ of the underlying graph Γ .

Boolean pairs. Consider a cube subgraph B of Γ , and consider the digraph \vec{B} of B oriented by o . One can easily see from the definition of an admissible orientation that \vec{B} is isomorphic to the Hasse diagram of a Boolean lattice. Hence \vec{B} determines the maximum element and the minimum element of the corresponding Boolean lattice. A pair (p, q) of vertices is called *o-Boolean*, or simply, *Boolean* if p and q are the minimum element and the maximum element, respectively, of the Boolean lattice associated with some cube subgraph of Γ . By convention, (p, p) is defined to be Boolean. The set of Boolean pairs is denoted by $\mathcal{B}(\mathbf{\Gamma})$. In Figure 2 in the introduction, for example, (p', p) , (v, q) , (p', v) are Boolean, and (q, p') is not Boolean.

Recall that any admissible orientation is acyclic (Lemma 2.4). Let $\preceq (= \preceq_o)$ be the transitive closure of relation $\swarrow (= \swarrow_o)$ on V_Γ . Then V_Γ is regarded as a partially ordered set according to this relation; so $p \swarrow q$ implies $p \prec q$. For any Boolean pair (p, q) , necessarily $p \preceq q$ holds.

Proposition 4.1. *Let $\mathbf{\Gamma}$ be a modular complex. For $p, q \in V_\Gamma$ with $p \preceq q$, we have the following:*

- (1) $[p, q]$ is a modular lattice, is convex in Γ , and is equal to $I(p, q)$.
- (2) (p, q) is Boolean if and only if $[p, q]$ is a complemented modular lattice.

In particular we can check whether a given pair is Boolean in time polynomial in $|V_\Gamma|$.

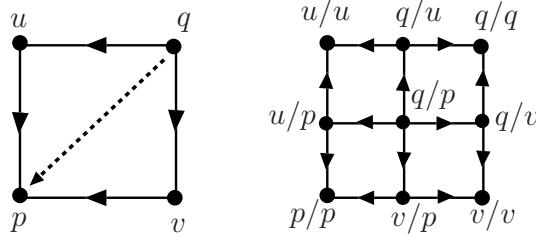


Figure 8: Orientations o and o^*

We prove this proposition in Section 4.3.1. We define the relation $\sqsubseteq (= \sqsubseteq_o)$ as: $p \sqsubseteq q$ if (p, q) is a Boolean pair. This relation \sqsubseteq coarsens \preceq , and is not transitive in general. Since a complemented modular lattice is relatively complemented (Theorem 2.11), we have:

$$(4.1) \quad \text{If } p \sqsubseteq q \text{ and } p \preceq u \preceq v \preceq q, \text{ then } p \sqsubseteq u \sqsubseteq v \sqsubseteq q.$$

For a vertex p , define subsets $\mathcal{L}_p^+(\Gamma)$ and $\mathcal{L}_p^-(\Gamma)$ of vertices by

$$(4.2) \quad \mathcal{L}_p^+(\Gamma) := \{q \in V_\Gamma \mid p \sqsubseteq q\}, \quad \mathcal{L}_p^-(\Gamma) := \{q \in V_\Gamma \mid q \sqsubseteq p\}.$$

In the sequel, $\mathcal{L}_p^+(\Gamma)$ and $\mathcal{L}_p^-(\Gamma)$ are often denoted by \mathcal{L}_p^+ and \mathcal{L}_p^- , respectively. Regard \mathcal{L}_p^+ as a poset by the partial order \preceq , and regard \mathcal{L}_p^- as a poset by the reverse of \preceq .

Theorem 4.2. *Let Γ be a modular complex. For every vertex p , both \mathcal{L}_p^+ and \mathcal{L}_p^- are complemented modular semilattices, and convex in Γ .*

Theorem 4.2 will be proved in Section 4.3.1. Therefore Γ is a structure obtained by gluing modular lattices and semilattices. Moreover Γ gives rise to a simplicial complex $\Delta(\Gamma)$ as follows. For each Boolean pair (p, q) and each ascending path $(p = p_0, p_1, p_2, \dots, p_k = q)$ from p to q , fill a k -dimensional simplex as in Figure 2. Then we obtain a simplicial complex $\Delta(\Gamma)$, and we can define an analogue of Lovász extension for any function on V_Γ . We however do not use this complex $\Delta(\Gamma)$ in the sequel, although our argument is based on this geometric view. Instead of dealing with $\Delta(\Gamma)$, we use a graph-theoretic operation, the 2-subdivision Γ^* of Γ , which comes from the barycentric subdivision of $\Delta(\Gamma)$.

2-subdivision and neighborhood semilattices. The 2-subdivision Γ^* of Γ is constructed as follows. A Boolean pair $(p, q) \in \mathcal{B}(\Gamma)$ is denoted by q/p . The 2-subdivision Γ^* of Γ is a simple undirected graph on the set $\mathcal{B}(\Gamma)$ of all Boolean pairs with edges given as: q/p and q'/p' are adjacent if and only if $p = p'$ and $qq' \in E_\Gamma$ or $q = q'$ and $pp' \in E_\Gamma$. The orientation o^* for Γ^* is given as: $q/p \prec_{o^*} q'/p'$ if $p = p'$ and $q \prec_o q'$ or if $q = q'$ and $p' \prec_o p$. See Figure 8. In fact, Γ^* does not depend on the choice of an admissible orientation; see [10].

An edge joining q/p and q'/p (resp. q/p and q/p') is denoted by qq'/p (resp. q/pp'). A function h^* on E_{Γ^*} is defined as $h^*(qq'/p) := h(qq')/2$ and $h^*(q/pp') := h(pp')/2$. Let $\Gamma^* := (\Gamma^*, o^*, h^*)$, which is called the 2-subdivision of Γ .

Theorem 4.3. *For a modular complex Γ , the 2-subdivision Γ^* is also a modular complex.*

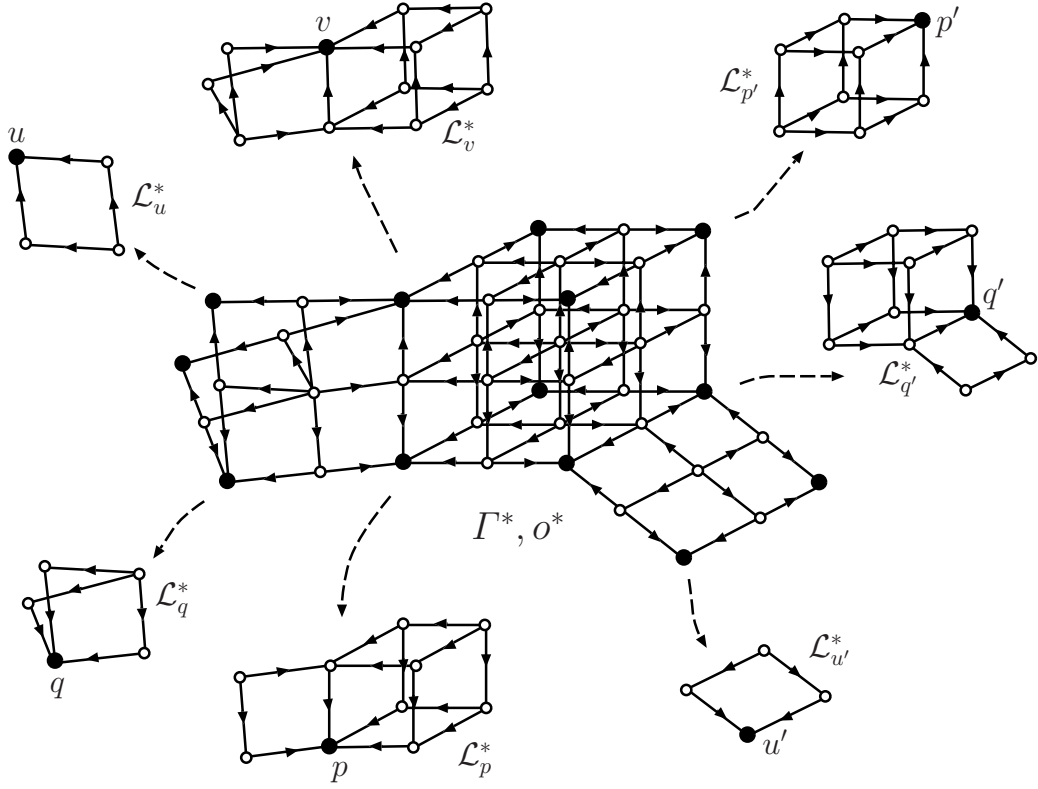


Figure 9: Construction of Γ^* and neighborhood semilattices

This theorem will be proved in Section 4.3.2. Figure 9 illustrates the 2-subdivision of Γ in Figure 2. By embedding $p \mapsto p/p$, we can regard $V_\Gamma \subseteq V_{\Gamma^*}$. The admissible orientation o^* is oriented so that the vertices in V_Γ are all sinks. The partial order \preceq_{o^*} on V_{Γ^*} induced by o^* is denoted by \preceq_* , and \sqsubseteq_{o^*} is also denoted by \sqsubseteq_* . In fact, one can show that two relations \preceq_* and \sqsubseteq_* are the same. Here we only note the following obvious relation:

$$(4.3) \quad q/p \preceq_* q'/p' \iff p' \preceq p \preceq q \preceq q' \quad (q/p, q'/p' \in \mathcal{B}(\Gamma)).$$

For each vertex $p \in V_\Gamma$, define the *neighborhood semilattice* $\mathcal{L}_p^* := \mathcal{L}_{p/p}^+(\Gamma^*)$. By Theorems 4.2 and 4.3, we obtain:

Proposition 4.4. *Let Γ be a modular complex. For each vertex p , the neighborhood semilattice \mathcal{L}_p^* is a complemented modular semilattice with the minimum element p .*

See Figure 9. Neighborhood semilattice \mathcal{L}_p^* has more information about the local property of p than \mathcal{L}_p^+ and \mathcal{L}_p^- have.

Valuation of local semilattices. A positive orbit-invariant function h naturally gives valuation v_p on \mathcal{L}_p^s for $s \in \{-, +\}$, and valuation v_p^* on \mathcal{L}_p^* by

$$(4.4) \quad \begin{aligned} v_p(q) &:= d_{\Gamma, h}(q, p) \quad (q \in \mathcal{L}_p^s, s \in \{-, +\}), \\ v_p^*(v/u) &:= d_{\Gamma^*, h^*}(v/u, p/p) \quad (v/u \in \mathcal{L}_p^*). \end{aligned}$$

See Lemma 2.14. In the sequel, semilattices \mathcal{L}_p^+ , \mathcal{L}_p^- , and \mathcal{L}_p^* are supposed to be endowed with these valuations.

Embedding of Γ into Γ^* . The distances on Γ and Γ^* are related as follows.

Proposition 4.5. *Let $\Gamma = (\Gamma, o, h)$ be a modular complex, and Γ^* the 2-subdivision of Γ . Then we have*

$$(4.5) \quad d_{\Gamma^*, h^*}(q/p, q'/p') = \frac{d_{\Gamma, h}(p, p') + d_{\Gamma, h}(q, q')}{2} \quad (q/p, q'/p' \in \mathcal{B}(\Gamma) = V_{\Gamma^*}).$$

In particular, $(V_\Gamma, d_{\Gamma, h})$ is isometrically embedded into $(V_{\Gamma^*}, d_{\Gamma^*, h^*})$ by $p \mapsto p/p$.

This proposition will be proved in Section 4.3.2.

Product of modular complexes. Suppose that we are given two modular complexes $\Gamma = (\Gamma, o, h)$ and $\Gamma' = (\Gamma', o', h')$. Then the Cartesian product $\Gamma \times \Gamma'$ is also modular. Furthermore, define the orientation $o \times o'$ of $\Gamma \times \Gamma'$ as: $(p, p') \swarrow_{o \times o'} (q, q')$ if $p' \swarrow_{o'} q'$ and $(p, p') \swarrow_{o \times o'} (q, p')$ if $p \swarrow_o q$. Then $o \times o'$ is an admissible orientation. Similarly define $h \times h'$ by $(h \times h')((p, p')(q, p')) := h(pq)$ and $(h \times h')((p, p')(p, q')) := h'(p'q')$, which is orbit-invariant in $\Gamma \times \Gamma'$. Thus we obtain a new modular complex $\Gamma \times \Gamma' := (\Gamma \times \Gamma', o \times o', h \times h')$, which is called the *product* of Γ and Γ' .

Lemma 4.6. *$(p, p') \sqsubseteq_{o \times o'} (q, q')$ if and only if $p \sqsubseteq_o q$ and $p' \sqsubseteq_{o'} q'$.*

Proof. Since $[(p, p'), (q, q')] \simeq [p, q] \times [p', q']$, $[(p, p'), (q, q')]$ is complemented modular if and only if both $[p, q]$ and $[p', q']$ are complemented modular. Thus, by Proposition 4.1, we have the claim. \square

In particular the correspondence $\mathcal{B}(\Gamma \times \Gamma') \ni (q, q')/(p, p') \mapsto (q/p, q'/p') \in \mathcal{B}(\Gamma) \times \mathcal{B}(\Gamma')$ is bijective, and we can regard

$$\mathcal{B}(\Gamma \times \Gamma') = \mathcal{B}(\Gamma) \times \mathcal{B}(\Gamma').$$

Under this correspondence, the product operation and the 2-subdivision operation commute in the following sense.

Lemma 4.7. (1) $\mathcal{L}_{(p, p')}^s(\Gamma \times \Gamma') = \mathcal{L}_p^s(\Gamma) \times \mathcal{L}_{p'}^s(\Gamma')$ for $s \in \{-, +\}$.

$$(2) (\Gamma \times \Gamma')^* = \Gamma^* \times \Gamma'^*.$$

$$(3) \mathcal{L}_{(p, p')}^*(\Gamma \times \Gamma') = \mathcal{L}_p^*(\Gamma) \times \mathcal{L}_{p'}^*(\Gamma').$$

Proof. (1) follows from the previous lemma. (2) follows from the fact that $(q, q')/(p, p')$ and $(v, v')/(u, u')$ have an edge in $(\Gamma \times \Gamma')^*$ if and only if $d_\Gamma(q, v) + d_{\Gamma'}(q', v') + d_\Gamma(p, u) + d_{\Gamma'}(p', u') = 1$, which is equivalent to the condition that $(q/p, q'/p')$ and $(v/u, v'/u')$ have an edge in $\Gamma^* \times \Gamma'^*$. (3) follows from (2). \square

4.2 L-convex function on modular complex

We are ready to introduce the concept of an L-convex function on a modular complex $\Gamma = (\Gamma, o, h)$. Consider the 2-subdivision Γ^* of Γ . For a function $g : V_\Gamma \rightarrow \mathbf{R}$, define $\bar{g} : V_{\Gamma^*} \rightarrow \mathbf{R}$ by

$$(4.6) \quad \bar{g}(q/p) := \frac{g(p) + g(q)}{2} \quad (q/p \in \mathcal{B}(\Gamma) = V_{\Gamma^*}).$$

This is the restriction of the Lovász extension of g ; see the introduction for the Lovász extension. By restricting \bar{g} to neighborhood semilattices, we obtain functions on (complemented) modular semilattices \mathcal{L}_p^* for each vertex p .

An *L-convex function* on Γ is a function $g : V_\Gamma \rightarrow \mathbf{R}$ such that for each vertex $p \in V_\Gamma$ the restriction of \bar{g} to \mathcal{L}_p^* is submodular on \mathcal{L}_p^* (with respect to the valuation v_p^*). Corresponding to Theorem 3.6, the distance function $d_{\Gamma, h}$ is L-convex on $\Gamma \times \Gamma$, which is one of the most important properties for our application to 0-extension problem.

Theorem 4.8. *For a modular complex $\Gamma = (\Gamma, o, h)$, the distance function $d_{\Gamma, h}$ is an L-convex function on $\Gamma \times \Gamma$.*

The proof of this theorem is given in Section 4.3.3. Corresponding to Lemma 3.7, we obtain:

Lemma 4.9. *Let Γ, Γ', Λ , and Λ' be modular complexes, and let g and g' be L-convex functions on Γ .*

(1) *For $b \in \mathbf{R}$ and $c, c' \in \mathbf{R}_+$, $b + cg + c'g'$ is an L-convex function on Γ .*

(2) *A function \tilde{g} defined by*

$$\tilde{g}(p, p') := g(p) \quad ((p, p') \in V_{\Gamma \times \Gamma'})$$

is an L-convex function on $\Gamma \times \Gamma'$.

(3) *Suppose that $\Gamma = \Lambda \times \Lambda'$. For $p' \in V_{\Lambda'}$, a function $g_{p'}$ defined by*

$$g_{p'}(p) := g(p, p') \quad (p \in V_\Lambda)$$

is an L-convex function on Λ .

Proof. (1) follows from Lemma 3.7 (1).

(2). $\overline{(\tilde{g})}(q/p, q'/p') = \{\tilde{g}(p, p') + \tilde{g}(q, q')\}/2 = \{g(p) + g(q)\}/2 = \overline{g}(q/p)$. Therefore, by Lemma 3.7 (2), function $(q/p, q'/p') \mapsto \overline{g}(q/p)$ is submodular on $\mathcal{L}_p^*(\Gamma) \times \mathcal{L}_{p'}^*(\Gamma') = \mathcal{L}_{(p,p')}^*(\Gamma \times \Gamma')$.

(3). $\overline{g_{p'}}(q/p) = \{g(p, p') + g(q, p')\}/2 = \overline{g}(q/p, p'/p')$. Since \overline{g} is submodular on $\mathcal{L}_p^*(\mathbf{A}) \times \mathcal{L}_{p'}^*(\mathbf{A}')$, by Lemma 3.7 (3), $\overline{g_{p'}}$ is submodular on $\mathcal{L}_p^*(\mathbf{A})$. \square

The restrictions of an L-convex function to \mathcal{L}_p^+ and to \mathcal{L}_p^- are submodular.

Lemma 4.10. *An L-convex function g on Γ is submodular on \mathcal{L}_p^+ and on \mathcal{L}_p^- for each vertex p .*

Proof. Consider the set \mathcal{L}_p^{*+} of elements in \mathcal{L}_p^* of the form q/p . By Proposition 4.5, any vertex in any shortest path between q/p and q'/p is of the form u/p . Hence \mathcal{L}_p^{*+} is convex in Γ and in \mathcal{L}_p^* . By Lemma 3.7 (4), \overline{g} is submodular on \mathcal{L}_p^{*+} . Obviously \mathcal{L}_p^{*+} is isomorphic to \mathcal{L}_p^+ by $q/p \mapsto q$. By using relation $g(q) = 2\overline{g}(q/p) - g(p)$ ($q \in \mathcal{L}_p^+$), we see the submodularity of g on \mathcal{L}_p^+ . \square

L-optimality criterion. Consider minimization of L-convex functions on a modular complex Γ . There is an optimality criterion that extends the L-optimality criterion of \mathbb{L}^1 -convex function in discrete convex analysis; see [47, Theorem 7.14].

Theorem 4.11 (L-optimality criterion). *Let g be an L-convex function on a modular complex Γ . For a vertex $p \in V_\Gamma$, the following conditions are equivalent:*

(1) $g(p) \leq g(q)$ holds for every $q \in V_\Gamma$.

(2) $g(p) \leq g(q)$ holds for every $q \in V_\Gamma$ with $p \sqsubseteq q$ or $q \sqsubseteq p$. That is

$$g(p) = \min\{g(q) \mid q \in \mathcal{L}_p^+\} = \min\{g(q) \mid q \in \mathcal{L}_p^-\}.$$

We prove this theorem in Section 4.3.4. The condition (2) implies that g can be minimized by tracing the 1-skeleton graph of $\Delta(\Gamma)$. By Lemma 4.10, checking the condition (2) reduces to the submodular function minimization on modular semilattices, analogous to the case of \mathbb{L}^1 -convex function in discrete convex analysis [47, Section 10.3].

Suppose that Γ is the product of modular complexes Γ_i for $i = 1, 2, \dots, n$. Again we say nothing about the complexity of the minimization under the oracle model. So we consider the VSCP situation. Recall Section 2.1. By an L-convex constraint on Γ we mean an L-convex function g on $\Gamma_{i_1} \times \Gamma_{i_2} \times \dots \times \Gamma_{i_k}$ for some $i_1 < i_2 < \dots < i_k$. By Lemma 4.9, the sum of L-convex constraints \mathcal{G} is an L-convex function on Γ . By L-optimality criterion, the optimality check of given vertex p reduces to the minimization of the sum of submodular constraints over modular semilattices $\mathcal{L}_p^s(\Gamma) = \mathcal{L}_{p_1}^s \times \mathcal{L}_{p_2}^s \times \dots \times \mathcal{L}_{p_n}^s$ for $s \in \{-, +\}$. By Theorem 3.9, we obtain:

Theorem 4.12. *Let Γ be the product of modular complexes Γ_i ($i = 1, 2, \dots, n$), and let \mathcal{G} be a set of L-convex constraints on Γ . Define $g : V_\Gamma \rightarrow \mathbf{R}$ by*

$$g(p) := \sum_{h \in \mathcal{G}} h(p_{I_h}) \quad (p = (p_1, p_2, \dots, p_n) \in V_\Gamma).$$

For a given vertex p , there exists an algorithm, in time polynomial in n and $|\mathcal{G}|L^K$, to find $q \in \mathcal{L}_p^- \cup \mathcal{L}_p^+$ with $g(q) < g(p)$ or conclude that p is a global minimizer of g , where $L := \max_{1 \leq i \leq n, s \in \{+, -\}} |\mathcal{L}_{p_i}^s|$ and $K := \max_{g \in \mathcal{G}} k_g$.

Steepest descent algorithm. Theorem 4.11, Lemma 4.10, and Theorem 3.9 naturally lead us to a descent algorithm for L-convex functions on modular complexes, analogous to the steepest descent algorithm for L-convex function minimization in discrete convex analysis.

Starting from an arbitrary point p , each descent step is to find, for $s \in \{-, +\}$, an optimal solution q^s of the problem:

$$(4.7) \quad \text{Minimize } g(q) \text{ over } q \in \mathcal{L}_p^s.$$

As mentioned already, this is a submodular function minimization. If $g(p) = g(q^+) = g(q^-)$, then p is optimal. Otherwise, take $s \in \{-, +\}$ with $g(q^s) = \min\{g(q^-), g(q^+)\} (< g(p))$, let $p := q^s$ (*steepest direction*), and repeat the descent step. After a finite number of descent steps, we can obtain an optimal solution (a minimizer of g).

In the case where f is an L^\natural -convex function on a box subset B of \mathbf{Z}^n , Murota [46] proved that, by appropriate choices of steepest directions, the number of the descent steps is bounded by l_1 -diameter of B ; later Kolmogorov and Shioura [40] improved this bound. We do not know whether a similar upper bound exists for L-convex function minimizations on general modular complexes. This issue will be studied in [26].

4.3 Proofs

4.3.1 Proof of Proposition 4.1 and Theorem 4.2

We can assume that h is the uniform unit edge weight, and $d_{\Gamma, h} = d_\Gamma$ is denoted by d . A path $(p_0, p_1, p_2, \dots, p_k)$ is said to be *ascending* if $p_i \swarrow p_{i+1}$ for $i = 0, \dots, k-1$.

Lemma 4.13. *For $p, q \in V_\Gamma$ with $p \preceq q$, a (p, q) -path P is shortest if and only if P is an ascending path from p to q . In particular, $I(p, q) = [p, q]$, any maximal chain in $[p, q]$ has the same length, and the rank r of $[p, q]$ is given by $r(a) = d(a, p)$.*

Proof. Suppose $p \preceq q$. Take an ascending path $P = (p = p_0, p_1, p_2, \dots, p_k = q)$. We use the induction on length k ; the statement for $k = 1$ is obvious.

(If part). We show $d(p, q) = k$. Suppose for contradiction that $d(p, q) < k$. By induction and bipartiteness, we have $d(p, p_k) = d(p, p_{k-1}) - 1$ and $d(p, p_{k-1}) = k - 1$. By the quadrangle condition (Lemma 2.3 (2)) for p_{k-1}, p_k, p_{k-2}, p , there is a common neighbor q^* of p_k, p_{k-2} with $d(p, q^*) = d(p, p_{k-1}) - 2 = k - 3$. Consider the 4-cycle of $p_{k-1}, p_k, q^*, p_{k-2}$. By orientability $p_{k-2} \swarrow q^*$ must hold. Hence we obtain an ascending path $(p = p_0, p_1, \dots, p_{k-2}, q^*)$ of length $k - 1$ with $d(p, q^*) = k - 3$. A contradiction.

(Only if part). Take any shortest path $Q = (p = q_0, q_1, q_2, \dots, q_{k'} = q)$ between p and q . We show that Q is ascending. By the if part, necessarily $k' = k$. It suffices to show that $p \preceq q_{k-1} \swarrow q$; by induction $(p = q_0, q_1, q_2, \dots, q_{k-1})$ is ascending, and hence Q is ascending. We can assume that $p_{k-1} \neq q_{k-1}$. By the quadrangle condition for q, p_{k-1}, q_{k-1}, p , there is a common neighbor q^* of p_{k-1}, q_{k-1} with $d(q^*, p) = d(q, p) - 2$. Then $q^* \in I(p, p_{k-1}) = [p, p_{k-1}]$ (by induction). This means $p \preceq q^* \swarrow p_{k-1}$, which in turn implies $q^* \swarrow q_{k-1} \swarrow q$ by the orientation of 4-cycle $(p_{k-1}, q, q_{k-1}, q^*)$. Thus $p \preceq q_{k-1} \swarrow q$, as required. \square

Let $(p)^\uparrow$ and $(p)^\downarrow$ denote the principal filter $\{q \in V_\Gamma \mid q \succeq p\}$ and the principal ideal $\{q \in V_\Gamma \mid q \preceq p\}$ of p , respectively.

Lemma 4.14. *For $a, b \in (p)^\uparrow$, there uniquely exists a median m of p, a, b , which coincides with $a \wedge b$. Similarly, for $a, b \in (p)^\downarrow$ there uniquely exists a median m of q, a, b , which coincides with $a \vee b$. Both $(p)^\uparrow$ and $(p)^\downarrow$ are convex, and modular semilattices.*

Proof. It suffices to prove the claim for $(p)^\uparrow$. Suppose that a, b, p have two distinct medians c, c' . Take a median m of c, c', p . Let $k := d(c, m) = d(c', m) > 0$. We can take an ascending path $(m = m_0, m_1, \dots, m_k = c)$ from m to c , and also can take a neighbor m' of m with $d(m, c') = 1 + d(m', c')$; necessarily $m \searrow m'$. By the quadrangle condition for m, m_1, m', a , there is a common neighbor m'_1 of m_1, m' such that $d(a, m'_1) = d(a, m) - 2$. Also by the quadrangle condition for m, m_1, m', b there is a common neighbor m''_1 of m_1, m' such that $d(b, m''_1) = d(b, m) - 2$. By $m_1 \searrow m \searrow m'$ and the orientability, we have $m_1 \searrow m'_1 \searrow m'$ and $m_1 \searrow m''_1 \searrow m'$. Hence $m'_1 = m''_1$ must hold. Similarly, by the quadrangle condition for m_1, m_2, m'_1, a and for m_1, m_2, m'_1, b , we can find a common neighbor m'_2 of m_2, m'_1 such that $d(m_2, a) = d(m'_2, a) + 1$ and $d(m_2, b) = d(m'_2, b) + 1$. Necessarily $m_2 \preceq m'_2 \preceq a, b$. Repeat this process to obtain a neighbor m'_k of $m_k (= c)$ such that $d(c, a) = d(m'_k, a) + 1$ and $d(c, b) = d(m'_k, b) + 1$. This implies that $d(a, b) \leq d(a, m'_k) + d(m'_k, b) = d(a, c) + d(c, b) - 2 = d(a, b) - 2$; a contradiction.

We show $m = a \wedge b$ in the poset $(p)^\uparrow$. Indeed, take an arbitrary $p' \in (p)^\uparrow$ with $a \succeq p' \preceq b$. Consider a median m' of a, b, p' . Since there is an ascending path from p to m' using p' , m' is also a median of a, b, p , and $m' = m$ by the uniqueness. Hence $p' \preceq m$.

We next show the convexity of $(p)^\uparrow$ by verifying (3) in Lemma 2.9. Clearly the subgraph of Γ induced by $(p)^\uparrow$ is connected. Take $a, b \in (p)^\uparrow$ with $d(a, b) = 2$. We show that $I(a, b) \subseteq (p)^\uparrow$. From Lemma 4.13, this is obvious when $a \preceq b$ or $b \preceq a$. Thus we may assume $a \not\preceq b$ and $b \not\preceq a$. Consider $a \wedge b$ in $(p)^\uparrow$ (the existence of $a \wedge b$ is guaranteed as above). By $d(a, b) = 2$, a and b cover $a \wedge b$. By the admissibility of the orientation o , if a and b have another common neighbor $c \neq a \wedge b$, then c covers a and b , i.e., $c = a \vee b \in (p)^\uparrow$. Hence $I(a, b) \subseteq (p)^\uparrow$, and $(p)^\uparrow$ is convex.

For arbitrary $a, b \in (p)^\uparrow$, every maximal common lower bound m' of a, b (not necessarily in $(p)^\uparrow$) belong to $I(a, b)$; consider a median m' of a, b , which is equal to m' by the above argument. By the convexity, m' belongs to $(p)^\uparrow$, and is equal to $a \wedge b$ in $(p)^\uparrow$. This means that a, b have the meet in the poset (V_Γ, \preceq) .

Now $(p)^\uparrow$ is a semilattice, and convex in Γ . The covering graph of $(p)^\uparrow$ is equal to the subgraph $\Gamma[(p)^\uparrow]$ induced by $(p)^\uparrow$, which is modular (by convexity). By Theorem 2.12, $(p)^\uparrow$ is a modular semilattice. \square

Proof of Proposition 4.1. (1). Notice that the convexity is closed under the intersection. By $[p, q] = (p)^\uparrow \cap (q)^\downarrow$ and by the previous lemma, we have the convexity of $[p, q]$. Also $[p, q]$ is an interval of a modular semilattice, and hence a modular lattice.

(2). If (p, q) is Boolean, then q is the join of atoms in $[p, q]$, and $[p, q]$ is a complemented modular lattice (Theorem 2.11). Conversely, if $[p, q]$ is a complemented modular lattice, then there is a Boolean sublattice of full rank (generated by a base), and (p, q) is Boolean.

We can check whether a pair (p, q) is Boolean by the following procedure. First construct the partial order \preceq , and the poset (V_Γ, \preceq) . If $p \not\preceq q$, then (p, q) is not Boolean. Suppose $p \preceq q$. Construct (or enumerate) $[p, q]$. For each $u \in [p, q]$, check the existence of an element (complement) $v \in [p, q]$ with $u \wedge v = p$ and $u \vee v = q$. If every element has a complement, then $[p, q]$ is complemented modular, and (p, q) is Boolean. Otherwise, (p, q) is not Boolean. This procedure can be done in time polynomial in $|V_\Gamma|$. \square

Proof of Theorem 4.2. It suffices to consider only \mathcal{L}_p^+ . The statement that \mathcal{L}_p^+ is a semilattice immediately follows from Lemma 4.14 and (4.1). Next we show the convexity. In view of Lemma 2.9, take $a, b \in \mathcal{L}_p^+$ with $d(a, b) = 2$, and take any common neighbor c of a, b . We show $c \in \mathcal{L}_p^+$. This is obvious if $a \searrow c \searrow b$ or $b \searrow c \searrow a$. Also, if $a \searrow c \searrow b$,

then $c = a \wedge b \in \mathcal{L}_p^+$ (if $c \neq a \wedge b$, then 4-cycle $(a, c, b, a \wedge b)$ violates the admissibility of the orientation o).

So suppose that $a \not\leq c \not\leq b$. By Lemma 4.13, both $[p, a]$ and $[p, b]$ belong to $[p, c]$. In particular, c is the join of a and b in $[p, c]$. Here a is the join of atoms in $[p, a]$ and b is the join of atoms in $[p, b]$. This means that c is the join of atoms in $[p, c]$. Hence modular lattice $[p, c]$ is complemented (Theorem 2.11), and hence (p, c) is Boolean.

The subgraph of Γ induced by any convex set is again a modular graph. Therefore the covering graph of \mathcal{L}_p^+ is modular. By Theorem 2.12, \mathcal{L}_p^+ is a modular semilattice. In particular, each $[p, q]$ for each $q \in \mathcal{L}_p^+$ is a complemented modular lattice, and \mathcal{L}_p^+ is a complemented modular semilattice. \square

4.3.2 Proof of Theorem 4.3 and Proposition 4.5

We start with preliminary results. By Proposition 4.1, interval $[p, q]$ is a modular lattice, and is convex in Γ . So we can consider the projection $\text{Pr}_{[p, q]} : V_\Gamma \rightarrow [p, q]$ (see Section 2.2.2).

Lemma 4.15. *For $p, q, p', q' \in V_\Gamma$ with $p \preceq q$ and $p' \preceq q'$, let u, v, u', v' be defined by*

$$(4.8) \quad u := \text{Pr}_{[p, q]}(p'), \quad v := \text{Pr}_{[p, q]}(q'), \quad u' := \text{Pr}_{[p', q']}(p), \quad v' := \text{Pr}_{[p', q']}(q).$$

Then we have:

- (1) $u \preceq v, u' \preceq v', \text{Pr}_{[p', q']}([p, q]) = [u', v']$, and $\text{Pr}_{[p, q]}([p', q']) = [u, v]$.
- (2) $[u, v]$ is isomorphic to $[u', v']$ by map $w \mapsto \text{Pr}_{[p', q']}(w)$. In particular, $p \sqsubseteq q$ implies $u' \sqsubseteq v'$.

Proof. The image of a convex set by the projection is again convex (Theorem 2.8), and a convex set in a modular lattice is exactly an interval (Lemma 2.16 (2)). Hence $\text{Pr}_{[p', q']}([p, q])$ is equal to a subinterval $[a', b']$ of $[p', q']$, and $\text{Pr}_{[p, q]}([p', q'])$ is equal to a subinterval $[a, b]$ of $[p, q]$. By Theorem 2.8, $\text{Pr}_{[p, q]}$ is an isometry between $[a, b]$ and $[a', b']$, and consequently this induces a graph isomorphism between the subgraphs induced by $[a, b]$ and $[a', b']$. By Lemma 2.5, for $g, h \in [a, b]$, if $g \not\leq h$ then $\text{Pr}_{[p', q']}(g) \not\leq \text{Pr}_{[p', q']}(h)$. Therefore, a poset $[a, b]$ is isomorphic to $[a', b']$ by $\text{Pr}_{[p', q']}$. Necessarily $\text{Pr}_{[p, q]}(a') = a$. Notice that a is the gate of p at $[a, b]$ (by Lemma 4.13). By Theorem 2.8 (3), $a = \text{Pr}_{[a, b]}(p) = \text{Pr}_{[p, q]} \circ \text{Pr}_{[p', q']}(p) = \text{Pr}_{[p, q]}(u')$. Then $\text{Pr}_{[p, q]}(u') = a = \text{Pr}_{[p, q]}(a')$ implies $a' = u'$ (by Theorem 2.8 (1)). Similarly $b' = v', a = u$, and $b = v$. Thus we obtain (1) and (2). \square

We use the same notation d for d_Γ and $d_{\Gamma^*, 1/2}$ (since they can be distinguished by the arguments).

Lemma 4.16. $d(q/p, q'/p') = \frac{d(p, p') + d(q, q')}{2}$ for $q/p, q'/p' \in V_{\Gamma^*}$.

Proof. Take a path $P = (q/p = q_0/p_0, q_1/p_1, \dots, q_k/p_k = q'/p')$ between q/p and q'/p' in Γ^* . The length of P is equal to $\sum_{i=0}^{k-1} (d(q_i, q_{i+1}) + d(p_i, p_{i+1}))/2 \geq (d(q, q') + d(p, p'))/2$; hence (\geq) holds.

We show the equality by the induction on $d(p, p') + d(q, q')$. Define u, v, u', v' by (4.8). Then $d(p, p') = d(p, u) + d(u, u') + d(u', p')$ and $d(q, q') = d(q, v) + d(v, v') + d(v', q')$ hold. Suppose $p \prec u$. Take an atom a of $[p, u]$. Then (a, q) is Boolean by (4.1), and q/a is adjacent to q/p in Γ^* . Also $d(a, u) = d(p, u) - 1$ (Lemma 4.13), implying

$d(a, p') = d(p, p') - 1$. By induction, $d(q/a, q'/p') = (d(a, p') + d(q, q'))/2$, and hence $d(q/p, q'/p') \leq (d(a, p') + d(q, q'))/2 + 1/2 = (d(p, p') + d(q, q'))/2$, as required.

Consequently we can assume $(p, q, p', q') = (u, v, u', v')$. In particular, $d(p, q) = d(p', q')$, $d(p, p') = d(q, q')$, and $d(p, q') = d(p, q) + d(q, q') = d(q, p')$ must hold. Take a neighbor a of p in $I(p, p')$; then $a \notin [p, q]$. Take a median b of q, a, q' . Then b must be a neighbor of q and satisfies $d(p, q) = d(a, b) = d(p, b) - 1 = d(a, q) - 1$. If $q \not\prec b$, then $p \not\prec a \preceq b$ (Lemma 4.13), and the join $q \vee a$ in $[p, b]$ is equal to b ; this means that $[p, b]$ is a complemented modular lattice, (p, b) is Boolean, and b/p is adjacent to q/p . Applying the induction to b/p and q'/p' , we obtain the equality $(=)$. Similarly, if $b \not\prec q$, then (a, q) is Boolean and apply the induction to q/a and q'/p' . \square

Proof of Theorem 4.3. Any 4-cycle in Γ^* is represented as $(q/p, q'/p, q'/p', q/p')$ for some edges pp', qq' in Γ , or $(x/p, y/p, z/p, w/p)$ or $(p/x, p/y, p/z, p/w)$ for 4-cycle (x, y, z, w) and vertex p in Γ . This immediately implies that o^* is an admissible orientation and h^* is orbit-invariant.

To show that Γ^* is modular, we are going to verify that Γ^* satisfies the two conditions of Lemma 2.3. If q/p and q'/p' are joined by an edge, then $d_\Gamma(p, q)$ and $d_\Gamma(p', q')$ have different parity. This implies that Γ^* is bipartite.

We next verify the quadrangle condition (Lemma 2.3 (2)). Take boolean pairs q/p and q'/p' . Suppose further that we are given two neighbors q_1/p_1 and q_2/p_2 of q/p with $d(q/p, q'/p') = 1/2 + d(q_1/p_1, q'/p') = 1/2 + d(q_2/p_2, q'/p')$. Our goal is to show the existence of a common neighbor q^*/p^* of $q_1/p_1, q_2/p_2$ with $d(q/p, q'/p') = 1 + d(q^*/p^*, q'/p')$.

It suffices to consider the following three cases:

- (i) $p_1 = p = p_2$.
- (ii) $p_1 = p$, $q_1 \not\prec q$, and $q_2 = q$.
- (iii) $p_1 \not\prec p = p_2$ and $q_1 = q \not\prec q_2$.

Case (i). By Lemma 4.16, we have $d(q, q') = 1 + d(q_i, q')$ for $i = 1, 2$. By Lemma 2.3 (2) for Γ , there is a common neighbor q^* of q_1, q_2 with $d(q, q') = 2 + d(q^*, q')$. Here $p \sqsubseteq q_i$ ($i = 1, 2$), and hence $q_i \in \mathcal{L}_p^+$. By the convexity of \mathcal{L}_p^+ (Theorem 4.2), we have $q^* \in \mathcal{L}_p^+$, implying $p \sqsubseteq q^*$. Again, by Lemma 4.16, we have $d(q/p, q'/p') = 1 + d(q^*/p, q'/p')$, as required.

Case (ii). We show $p_2 \sqsubseteq q_1$, which implies that q_1/p_2 is a required common neighbor (by Lemma 4.16). If $p_2 \not\prec p$, then $p_2 \sqsubseteq q$ and $p_2 \not\prec p \preceq q_1 \not\prec q$ imply $p_2 \sqsubseteq q_1$ (by (4.1)). Suppose $p \not\prec p_2$. By Lemma 4.15, $\text{Pr}_{[p, q]}([p', q'])$ is equal to interval $[a, b]$ for $a = \text{Pr}_{[p, q]}(p')$ and $b = \text{Pr}_{[p, q]}(q')$. Then $d(p_2, p') = d(p, p') - 1$ implies $p_2 \preceq a$. Similarly $b \preceq q_1$. Thus $p \preceq p_2 \preceq a \preceq b \preceq q_1 \preceq q$ and $p \sqsubseteq q$ imply $p_2 \sqsubseteq q_1$ (by (4.1)), as required.

Case (iii). We show $p_1 \sqsubseteq q_2$; then q_2/p_1 is a common neighbor as required. We use the induction on $d(p_1, q_2)$; in the case of $d(p_1, q_2) = 2$ the subsequent argument shows that $I(p_1, q_2) = [p_1, q_2]$ contains an element different from p_1, q_2, p , and hence (p_1, q_2) is Boolean. Since $[p_1, q_2]$ is a convex set, we can consider $\text{Pr}_{[p_1, q_2]}([p', q'])$, which is equal to $[u, v]$ for $u := \text{Pr}_{[p_1, q_2]}(p')$, $v := \text{Pr}_{[p_1, q_2]}(q')$ (Lemma 4.15). Then necessarily $p \not\prec u$; otherwise $d(p_1, p') = d(p_1, u) + d(u, p') = 1 + d(p, u) + d(u, p') = 1 + d(p, p')$, contradicting the first assumption $d(p_1, p') = d(p, p') - 1$.

Suppose $p_1 \neq u$. Take an atom $a \in [p_1, u]$. If $a \not\prec q$, then q_2 is equal to the join $a \vee q$, and consequently the join of atoms in $[p_1, q_2]$. Thus $[p_1, q_2]$ is complemented, and $p_1 \sqsubseteq q_2$ (Proposition 4.1). Suppose that $a \preceq q$. Consider the join $a \vee p$ in $[p_1, q_2]$, which belongs

to $[p, q]$. Then $d(a \vee p, p') = d(a \vee p, u) + d(u, p') = 1 + d(a, u) + d(u, p') = 1 + d(a, p')$ (by $a \vee p \not\leq u$ and Lemma 4.13). By (4.1), both (a, q_1) and $(a \vee p, q_2)$ are Boolean. By induction, $a \sqsubseteq q_2$, and $[a, q_2]$ is complemented modular. Thus we can take an atom $b (\neq a \vee p)$ of $[a, q_2]$ with $q \vee b = q_2$. Necessarily $p \not\leq b$ (otherwise both a and p are covered by c and $a \vee p$; contradicting the fact that $[p_1, q_2]$ is a lattice). By modularity equality with $p \wedge b = p_1$, the join $p \vee b$ has rank 3 in $[p_1, q_2]$. We can take c with $p \not\leq c \not\leq b \vee p$. Then $d(p, b) = 3$, and $d(p_1, b) = d(c, b) = 2$. By the quadrangle condition, there is a common neighbor w of p_1, c, b , which is an atom of $[p_1, q]$. Then $a \vee w = b$, and $q_2 = q \vee b = q \vee a \vee w$. This means that q_2 is the join of atoms of $[p_1, q_2]$; thus $p_1 \sqsubseteq q_2$.

By the same argument (for v, q_2), we can assume that $u = p_1$ and $v = q_2$. Since $[u, v]$ is isomorphic to a subinterval of complemented modular lattice $[p', q']$ (Lemma 4.15), $[p_1, q_2] (= [u, v])$ is also a complemented modular lattice, implying $p_1 \sqsubseteq q_2$. \square

Proof of Proposition 4.5. We have proved (4.5) for the case $h = 1$ in Lemma 4.16. By Theorem 4.3 shown above, Γ^* is now a modular graph. Let P be a shortest path with respect to h^* . By Lemma 2.7, this is also shortest with respect to uniform edge-length $1/2$. Necessarily the paths obtained from $(q = q_0, q_1, q_2, \dots, q_k = q')$ and $(p = p_0, p_1, p_2, \dots, p_k = p')$ (by identifying repetitions) are both shortest in Γ with respect to uniform edge-length 1. Again, by Lemma 2.7, they are shortest relative to h , and have the lengths $d_{\Gamma, h}(p, p')$ and $d_{\Gamma, h}(q, q')$, respectively. Thus (4.5) holds. \square

4.3.3 Proof of Theorems 3.6 and 4.8

Consider the 2-subdivision $(\Gamma \times \Gamma)^*$, which is identified with $\Gamma^* \times \Gamma^*$ by correspondence $(q, q')/(p, p') \leftrightarrow (q/p, q'/p')$ (Lemma 4.7). Consider $\overline{d_{\Gamma, h}} : V_{(\Gamma \times \Gamma)^*} \rightarrow \mathbf{R}$. Then we have

$$\overline{d_{\Gamma, h}}((q, q')/(p, p')) = \frac{d_{\Gamma, h}(p, p') + d_{\Gamma, h}(q, q')}{2} = d_{\Gamma^*, h^*}(q/p, q'/p'),$$

where the first equality is the definition (4.6) and the second follows from Proposition 4.5. Hence it suffices to show that $d_{\Gamma^*, h^*} : V_{\Gamma^*} \times V_{\Gamma^*} \rightarrow \mathbf{R}$ is submodular on $\mathcal{L}_{(a/a, b/b)}^+(\Gamma^* \times \Gamma^*) = \mathcal{L}_{a/a}^+(\Gamma^*) \times \mathcal{L}_{b/b}^+(\Gamma^*)$ for every $(a, b) \in V_{\Gamma} \times V_{\Gamma}$. Therefore, by taking Γ^* as Γ , Theorem 4.8 follows from the following.

Lemma 4.17. *The distance function $d_{\Gamma, h}$ is submodular on $\mathcal{L}_a^+ \times \mathcal{L}_b^+$ for every $a, b \in V_{\Gamma}$.*

Proof. By Proposition 3.8, it suffices to show the following, where we denote $d_{\Gamma, h}$ by d , and denote the valuation on \mathcal{L}_a^+ (defined in (4.4)) by v .

- (1) For every $u \in \mathcal{L}_b^+$ and every antipodal pair (p, q) in \mathcal{L}_a^+ , we have

$$v[p \wedge q, q]d(p, u) + v[p \wedge q, p]d(q, u) \geq (v[p \wedge q, p] + v[p \wedge q, q])d(p \wedge q, u).$$

- (2) For every $u \in \mathcal{L}_b^+$ and every 2-bounded pair (p, q) in \mathcal{L}_a^+ , we have

$$d(p, u) + d(q, u) \geq d(p \wedge q, u) + d(p \vee q, u).$$

- (2') For every $p, q \in \mathcal{L}_a^+$ with $p \not\leq q$ and every $p', q' \in \mathcal{L}_b^+$ with $p' \not\leq q'$, we have

$$d(q, p') + d(p, q') \geq d(q, q') + d(p, p').$$

Note that (2) and (2') correspond to the submodularity condition for 2-bounded pairs.

(1). We may assume that $p \wedge q = a$ (by considering $\mathcal{L}_{p \wedge q}^+$) and $v(a) = 0$. Take a median m of p, q, u . By $m \in I(p, q)$ and Lemma 2.15, there are $p' \in [a, p]$ and $q' \in [a, q]$ with $m = p' \vee q'$. Let $D := d(m, u)$. Then we have $d(p, u) = v[p', p] + v(q') + D$, $d(q, u) = v[q', q] + v(p') + D$, and $d(a, u) \leq v(p') + v(q') + D$. Hence we get

$$\begin{aligned} & v(q)d(p, u) + v(p)d(q, u) - (v(p) + v(q))d(a, u) \\ & \geq v(q)\{v[p', p] + v(q') + D\} + v(p)\{v[q', q] + v(p') + D\} \\ & \quad - \{v(p) + v(q)\}\{v(p') + v(q') + D\} \\ & = v(q)v[p', p] + v(p)v[q', q] - v(p)v(q') - v(q)v(p') \\ & = 2v[q', q]v[p', p] - 2v(p')v(q'), \end{aligned}$$

where we use $v(p) = v(p') + v[p', p]$ and $v(q) = v(q') + v[q', q]$ for the last equality. This must be nonnegative since (p, q) is antipodal; see (3.9)

(2). Recall the notion of gated sets (Section 2.2.2); $[p \wedge q, p \vee q]$ is convex, and is gated (Lemmas 2.9 and 2.16). Let $m := \text{Pr}_{[p \wedge q, p \vee q]}(u)$, and $D := d(m, u)$. Then we have $d(x, u) = d(x, m) + D$ for $x \in \{p, q, p \wedge q, p \vee q\}$. There are three cases: (i) $m \in \{p, q\}$, (ii) $m \in \{p \wedge q, p \vee q\}$, and (iii) $m \notin \{p, q, p \wedge q, p \vee q\}$. Note that $(p, p \wedge q, q, p \vee q)$ forms a 4-cycle since (p, q) is 2-bounded. Let $\alpha := v[p, p \vee q] = v[p \wedge q, q]$ and $\beta := v[p \wedge q, p] = v[q, p \vee q]$. Consider the case (i). Then $\{d(p \wedge q, m), d(p \vee q, m)\} = \{\alpha, \beta\}$ and $\{d(p, m), d(q, m)\} = \{0, \alpha + \beta\}$. Hence $d(p, u) + d(q, u) - d(p \wedge q, u) - d(p \vee q, u) = 0$. Consider the case (ii). Then $\{d(p, m), d(q, m)\} = \{\alpha, \beta\}$, and $\{d(p \wedge q, m), d(p \vee q, m)\} = \{0, \alpha + \beta\}$. Hence $d(p, u) + d(q, u) - d(p \wedge q, u) - d(p \vee q, u) = 0$. Consider the case (iii). Then m is a common neighbor of $p \wedge q, p \vee q$ different from p, q . Hence all edges in $[p \wedge q, p \vee q]$ belong to the same orbit. Thus $\alpha = \beta$, $d(p, m) = d(q, m) = 2\alpha$, $d(p \wedge q, m) = d(p \vee q, m) = \alpha$, and $d(p, u) + d(q, u) - d(p \wedge q, u) - d(p \vee q, u) = 2\alpha > 0$.

(2'). Consider $\text{Pr}_{\{p', q'\}}(\{p, q\})$. Let $D := d(\{p, q\}, \{p', q'\})$. There are two cases: (i) $|\text{Pr}_{\{p', q'\}}(\{p, q\})| = 1$ and (ii) $\{p', q'\} = \text{Pr}_{\{p', q'\}}(\{p, q\})$. Consider the case (i). For u, v, u', v' with $\{u, v\} = \{p, q\}$ and $\{u', v'\} = \{p', q'\}$, we have $d(v, u') = D$, $d(u, u') = D + h(uv)$, $d(v, v') = D + h(u'v')$, and $d(u, v') = D + h(uv) + h(u'v')$. Thus $d(u, u') + d(v, v') = d(u, v') + d(v, u')$, and the equality holds in (2'). Consider the case (ii). By Theorem 2.8 and Lemma 2.5, we have $p' = \text{Pr}_{\{p', q'\}}(p)$, $q' = \text{Pr}_{\{p', q'\}}(q)$, $d(p, p') = d(q, q') = D$, and that pq and $p'q'$ must belong to the same orbit Q . Then $d(p, q') = d(q, p') = D + h_Q$. Therefore (2') holds. \square

The above proof works even when \mathcal{L}_a^+ and \mathcal{L}_b^+ are replaced by the principal filters $(a)^\uparrow$ and $(b)^\uparrow$, respectively, since they are convex and are (not necessarily complemented) modular semilattices (Lemma 4.14). Therefore Theorem 3.6 follows from Theorem 2.12 and:

Lemma 4.18. *The distance function $d_{\Gamma, h}$ is submodular on $(a)^\uparrow \times (b)^\uparrow$ for every $a, b \in V_\Gamma$.*

4.3.4 Proof of L-optimality criterion (Theorem 4.11)

Let $\mathbf{\Gamma} = (\Gamma, o, h)$ be a modular complex and let g be a function on V_Γ . Let $\bar{\Gamma}$ denote the graph obtained from Γ by joining all Boolean pairs (p, q) (with $d_\Gamma(p, q) \geq 2$). Namely $\bar{\Gamma}$ is the 1-skeleton graph of the complex $\Delta(\mathbf{\Gamma})$. For $\alpha \in \mathbf{R}$, the *level-set subgraph* $\bar{\Gamma}_{g, \alpha}$ is the subgraph of $\bar{\Gamma}$ induced by the set of vertices p with $g(p) \leq \alpha$. The following connectivity property of $\bar{\Gamma}_{g, \alpha}$ rephrases the L-optimality criterion (Theorem 4.11).

Proposition 4.19. *Let g be an L-convex function on $\mathbf{\Gamma}$, and let $l := \min_{p \in V_\Gamma} g(p)$. For every $\alpha \geq l$, the level-set graph $\bar{\Gamma}_{g, \alpha}$ is connected. In addition, if $\alpha > l$, then every vertex in $g^{-1}(\alpha)$ is adjacent to a vertex of $\bar{\Gamma}_{g, \alpha} \setminus g^{-1}(\alpha)$ in $\bar{\Gamma}_{g, \alpha}$.*

In a crucial step of the proof, we use the following general property of submodular functions on a modular semilattice, where a sequence (p_0, p_1, \dots, p_m) of elements in a poset is said to be *comparable* if $p_i \preceq p_{i+1}$ or $p_{i+1} \preceq p_i$ for $i = 0, 1, 2, \dots, m$.

Lemma 4.20. *Let f be a submodular function on a modular semilattice \mathcal{L} . For $p, q \in \mathcal{L}$ and $\alpha \in \mathbf{R}$, if $f(p) \leq \alpha$ and $f(q) < \alpha$, there exists a comparable sequence $(p = p_0, p_1, p_2, \dots, p_m = q)$ such that $f(p_i) < \alpha$ for $i = 1, 2, \dots, m$.*

Proof. We can assume that $\alpha = 0$ by letting $f \leftarrow f - \alpha$. Also we may assume that p and q are incomparable; $\text{Conv } I(p, q)$ is a polygon, and hence $[C(p; p, q)] < 1$ and $[C(q; p, q)] < 1$. Consider inequality (3.6):

$$(1 - [C(p; p, q)])f(p) + (1 - [C(q; p, q)])f(q) \geq f(p \wedge q) + \sum_{u \in \mathcal{E}(p, q) \setminus \{p, q\}} [C(u; p, q)]f(u).$$

Then the left hand side is negative, and hence the right hand side is negative. If $f(p \wedge q) < 0$, then $(p, p \wedge q, q)$ is a required sequence. Otherwise there exists $u \in \mathcal{E}(p, q) \setminus \{p, q\}$ with $f(u) < 0$ (since $[C(u; p, q)]$ is nonnegative). By an inductive argument (on distance between p and q), there are comparable sequences $(p, p_1, p_2, \dots, p_k = u)$ and $(u, q_1, q_2, \dots, q_{k'} = q)$ with $f(p_i) < 0$ and $f(q_j) < 0$. Concatenating them, we obtain a required sequence. \square

Proof of Proposition 4.19. Suppose (indirectly) that $\bar{\Gamma}_{g, \alpha'}$ is disconnected for some α' . For a sufficiently large α , the graph $\bar{\Gamma}_{g, \alpha}$ is equal to $\bar{\Gamma}$, and is connected (since Γ is finite and g has no infinite value). Also, for a sufficiently small $\epsilon > 0$, it holds $\bar{\Gamma}_{g, \alpha - \epsilon} = \bar{\Gamma}_{g, \alpha} \setminus g^{-1}(\alpha)$. This implies that there exists $\alpha \geq l$ such that $\bar{\Gamma}_{g, \alpha}$ is connected, and $\bar{\Gamma}_{g, \alpha} \setminus g^{-1}(\alpha)$ is disconnected. Then there exists a pair of vertices p, p' belonging to different components in $\bar{\Gamma}_{g, \alpha} \setminus g^{-1}(\alpha)$; in particular $g(p) < \alpha$ and $g(p') < \alpha$. Take such a pair (p, p') with $k := d_{\bar{\Gamma}_{g, \alpha}}(p, p')$ minimum. There exists a path $(p = p_0, p_1, \dots, p_k = p')$ in $\bar{\Gamma}_{g, \alpha}$ with $g(p_i) = \alpha$ for $i = 1, 2, \dots, k - 1$.

We first show $k = 2$. Consider $\mathcal{L}_{p_1}^*$ and \bar{g} on $\mathcal{L}_{p_1}^*$. We may assume that $p_1 \sqsubseteq p_2$. Let $u := p/p_1$ if $p_1 \sqsubseteq p$ and $u := p_1/p$ if $p \sqsubseteq p_1$. Then $\bar{g}(u) < \alpha$ and $\bar{g}(p_2/p_1) \leq \alpha$. Therefore, by Lemma 4.20, there exists a comparable sequence $(u = u_0, u_1, \dots, u_{m-1}, u_m = p_2/p_1)$ in $\mathcal{L}_{p_1}^*$ such that $\bar{g}(u_i) < \alpha$ for $i = 0, 1, 2, \dots, m - 1$. Consider u_{m-1} , which is equal to q'/q for some $q, q' \in V_\Gamma$ with $q \sqsubseteq q'$. Then (i) $p_2/p_1 \sqsubseteq_* q'/q$ or (ii) $q'/q \sqsubseteq_* p_2/p_1$.

Consider case (i). By (4.3), we have $q \preceq p_1 \preceq p_2 \preceq q'$. By (4.1), we have $q \sqsubseteq p_i \sqsubseteq q'$ for $i = 1, 2$. Thus both q and q' are adjacent to each of p_1 and p_2 (in $\bar{\Gamma}$). By $\bar{g}(u_{m-1}) < \alpha$, we have $g(q) < \alpha$ or $g(q') < \alpha$. Say $g(q) < \alpha$; q is adjacent to p_1 and p_2 in $\bar{\Gamma}_{g, \alpha}$. If q and p' belongs different components in $\bar{\Gamma} \setminus g^{-1}(\alpha)$, then path $(q, p_2, p_3, \dots, p_k = p')$ violates the minimality assumption. This means that q and p' belong to the same component, which is different from the component that p belongs to. Thus we could have chosen path (p, p_1, q) of length 2. This implies that $k = 2$ and $g(p_2) < \alpha$.

Consider case (ii). By (4.3) and $q'/q \in \mathcal{L}_{p_1}^*$, we have $p_1 \preceq q \preceq q' \preceq p_2$ and $q \preceq p_1 \preceq q'$. Hence $q = p_1$, and $p_1 \sqsubseteq q' \sqsubseteq p_2$ (by (4.1)). Also, we have $g(q') < \alpha$, and q' is adjacent to each of p_1 and p_2 . As above, by the minimality, we must have $k = 2$ and $g(p_2) < \alpha$.

Suppose that u_i is represented by $u_i = q'_i/q_i$ for $q_i, q'_i \in V_\Gamma$ with $q_i \sqsubseteq q'_i$ ($i = 0, 1, 2, \dots, m$). Then $q_i \preceq q_{i+1} \preceq q'_{i+1} \preceq q'_i$ or $q_{i+1} \preceq q_i \preceq q'_i \preceq q'_{i+1}$. Again by (4.1), both q_i and q'_i are adjacent to each of q_{i+1} and q'_{i+1} in $\bar{\Gamma}$. Also, by $\bar{g}(q'_i/q_i) < \alpha$, at least one of $g(q_i)$ and $g(q'_i)$ is less than α . This means that there is a path in $\bar{\Gamma}_{g, \alpha} \setminus g^{-1}(\alpha)$ connecting p and p' . This is a contradiction to the initial assumption that p and p' belong to distinct components in $\bar{\Gamma}_{g, \alpha} \setminus g^{-1}(\alpha)$.

We show the latter part. Take $p \in g^{-1}(\alpha)$. Then there is a pair of $q \in V_\Gamma$ and a path $(q = p_0, p_1, p_2, \dots, p_k = p)$ in $\bar{\Gamma}$ such that $g(q) < \alpha$ and $g(p_i) = \alpha$ for $i = 1, 2, \dots, k$. Take such a pair with the minimum length k . We show $k = 1$. Suppose that $k \geq 2$. As above, by considering \bar{g} on $\mathcal{L}_{p_1}^*$, we can find a neighbor q' of p_2 with $f(q') < \alpha$. This is a contradiction to the minimality of k . Hence $k = 1$, implying the latter statement. \square

5 Minimum 0-extension problems

In this section, we study, from the viewpoint developed in the previous sections, the minimum 0-extension problem **0-Ext** $[\Gamma]$ on an orientable modular graph Γ . In Section 5.1, we verify that **0-Ext** $[\Gamma]$ can be formulated as an L-convex function minimization on a modular complex $\mathbf{\Gamma}^n$. In Section 5.2, we present a powerful optimality criterion (Theorem 5.2) for **0-Ext** $[\Gamma]$ by specializing the L-optimality criterion (Theorem 4.11). In Section 5.3 we prove the main theorem (Theorem 1.6) of this paper. In Section 5.4 we consider the minimum 0-extension problem for metrics, not necessarily graph metrics, and extend Theorem 1.6 to metrics.

5.1 L-convexity of multifacility location functions

Let Γ be an orientable modular graph with an orbit-invariant function h . We are given a finite set V with $V_\Gamma \subseteq V$. Suppose that $V \setminus V_\Gamma = \{1, 2, \dots, n\}$. For a nonnegative cost $c : \binom{V}{2} \rightarrow \mathbf{Q}_+$, let $(c \cdot d_{\Gamma, h})$ be a function on $(V_\Gamma)^n$ defined by

$$(c \cdot d_{\Gamma, h})(\rho) := \sum_{s \in V_\Gamma} \sum_{1 \leq j \leq n} c(sj) d_{\Gamma, h}(s, \rho_j) + \sum_{1 \leq i < j \leq n} c(ij) d_{\Gamma, h}(\rho_i, \rho_j) \\ (\rho = (\rho_1, \rho_2, \dots, \rho_n) \in (V_\Gamma)^n).$$

Such a function is called a *multifacility location function* on Γ . A point ρ in $(V_\Gamma)^n$ is called a *location*. Consider the following natural weighted version of **0-Ext** $[\Gamma]$:

Multifac $[\Gamma, h; V, c]$: Minimize $(c \cdot d_{\Gamma, h})(\rho)$ over all locations $\rho \in (V_\Gamma)^n$,

where the unweighted version corresponds to $h = 1$.

Fix an admissible orientation o of Γ . By a natural identification $(V_\Gamma)^n \simeq V_{\Gamma^n}$, a location is regarded as a vertex in V_{Γ^n} . In particular, $(c \cdot d_{\Gamma, h})$ is regarded as a function on V_{Γ^n} . By Theorem 4.8 and Lemma 4.9, we have:

Theorem 5.1. *Multifacility location function $(c \cdot d_{\Gamma, h})$ is an L-convex function on modular complex $\mathbf{\Gamma}^n$.*

Therefore **Multifac** $[\Gamma, h; V, c]$ is an L-convex function minimization on modular complex $\mathbf{\Gamma}^n$. So we can apply the results in the previous section to **Multifac** $[\Gamma, h; V, c]$.

5.2 Optimality criterion and orbit-additivity

Let $\rho = (\rho_1, \rho_2, \dots, \rho_n)$ be a location. A location $\rho' = (\rho'_1, \rho'_2, \dots, \rho'_n)$ is said to be a *forward neighbor* of ρ if $\rho'_i \in \mathcal{L}_{\rho_i}^+$ for all i , and is said to be a *backward neighbor* of ρ if $\rho'_i \in \mathcal{L}_{\rho_i}^-$ for all i . A forward or backward neighbor is simply called a *neighbor*. This terminology is due to [22]. By Lemma 4.7, the set of forward (resp. backward) neighbors of ρ is equal to $\mathcal{L}_\rho^+(\mathbf{\Gamma}^n) = \mathcal{L}_{\rho_1}^+ \times \mathcal{L}_{\rho_2}^+ \times \dots \times \mathcal{L}_{\rho_n}^+$ (resp. $\mathcal{L}_\rho^-(\mathbf{\Gamma}^n) = \mathcal{L}_{\rho_1}^- \times \mathcal{L}_{\rho_2}^- \times \dots \times \mathcal{L}_{\rho_n}^-$).

We introduce a sharper concept of a neighbor using orbits. Recall (2.11) in Section 2.3 for definition of $\mathcal{L}|Q$. Let Q be an orbit of Γ ; consider $\mathcal{L}_\rho^\pm|Q := \mathcal{L}_\rho^\pm|(Q \cap$

{edges belonging to \mathcal{L}_ρ^\pm }). A forward neighbor ρ' of ρ is called a *forward Q -neighbor* of ρ if $\rho'_i \in \mathcal{L}_{\rho_i}^+|Q$ for all i , and is called a *backward Q -neighbor* of ρ if $\rho'_i \in \mathcal{L}_{\rho_i}^-|Q$ for all i . The set of all forward (reps. backward) Q -neighbors of ρ is denoted by $\mathcal{L}_\rho^+|Q = \mathcal{L}_{\rho_1}^+|Q \times \mathcal{L}_{\rho_2}^+|Q \times \cdots \times \mathcal{L}_{\rho_n}^+|Q$ (resp. $\mathcal{L}_\rho^-|Q = \mathcal{L}_{\rho_1}^-|Q \times \mathcal{L}_{\rho_2}^-|Q \times \cdots \times \mathcal{L}_{\rho_n}^-|Q$). A forward or backward Q -neighbor is simply called a Q -neighbor.

The main result in this section is the following optimality criterion, which has been shown for some special cases of orientable modular graphs: trees by Kolen [37, Chapter 3], median graphs by Chepoi [12, p.11–12], and frames by Hirai [22, Section 4.1].

Theorem 5.2. *Let Γ be an orientable modular graph with an admissible orientation o and a positive orbit-invariant function h . For a location ρ , the following conditions are equivalent:*

(1) ρ is optimal to **Multifac** $[\Gamma, h; V, c]$.

(2) ρ is optimal to **Multifac** $[\Gamma, 1; V, c]$.

(3) For every neighbor ρ' of ρ , we have $(c \cdot d_{\Gamma, h})(\rho) \leq (c \cdot d_{\Gamma, h})(\rho')$. That is

$$(c \cdot d_{\Gamma, h})(\rho) = \min\{(c \cdot d_{\Gamma, h})(\rho') \mid \rho' \in \mathcal{L}_\rho^+\} = \min\{(c \cdot d_{\Gamma, h})(\rho') \mid \rho' \in \mathcal{L}_\rho^-\}.$$

(4) For every neighbor ρ' of ρ , we have $(c \cdot d_{\Gamma, 1})(\rho) \leq (c \cdot d_{\Gamma, 1})(\rho')$. That is

$$(c \cdot d_{\Gamma, 1})(\rho) = \min\{(c \cdot d_{\Gamma, 1})(\rho') \mid \rho' \in \mathcal{L}_\rho^+\} = \min\{(c \cdot d_{\Gamma, 1})(\rho') \mid \rho' \in \mathcal{L}_\rho^-\}.$$

(5) For every orbit Q and every Q -neighbor ρ' of ρ , we have $(c \cdot d_{\Gamma, 1})(\rho) \leq (c \cdot d_{\Gamma, 1})(\rho')$. That is, for every orbit Q , we have

$$(c \cdot d_{\Gamma, 1})(\rho) = \min\{(c \cdot d_{\Gamma, 1})(\rho') \mid \rho' \in \mathcal{L}_\rho^+|Q\} = \min\{(c \cdot d_{\Gamma, 1})(\rho') \mid \rho' \in \mathcal{L}_\rho^-|Q\}.$$

Any orbit Q in Γ forms an orbit-union in the covering graph of each \mathcal{L}_ρ^+ . By Lemma 2.16 (3), $\mathcal{L}_\rho^+|Q$ is a complemented modular lattice, and convex in \mathcal{L}_ρ^+ . Thus all conditions (3), (4), and (5) are checked by submodular function minimization on modular semilattice.

Before the proof, we explain consequences of Theorem 5.2. The first consequence is that in solving **Multifac** $[\Gamma, h; V, c]$, we may replace h with the unit function, even when h is not positive.

Theorem 5.3. *For every nonnegative orbit-invariant function h , every optimal location in **Multifac** $[\Gamma, 1; V, c]$ is optimal to **Multifac** $[\Gamma, h; V, c]$.*

Proof. Let ρ be an optimal location for **Multifac** $[\Gamma, 1; V, c]$. Take an arbitrary positive $\epsilon > 0$. Consider the positive orbit invariant function $h + \epsilon 1$. By Theorem 5.2, ρ is optimal to **Multifac** $[\Gamma, h + \epsilon 1; V, c]$. Hence, for an arbitrary location ρ' , we have

$$(c \cdot d_{\Gamma, h})(\rho') + \epsilon(c \cdot d_{\Gamma, 1})(\rho') = (c \cdot d_{\Gamma, h + \epsilon 1})(\rho') \geq (c \cdot d_{\Gamma, h + \epsilon 1})(\rho) = (c \cdot d_{\Gamma, h})(\rho) + \epsilon(c \cdot d_{\Gamma, 1})(\rho).$$

Since $\epsilon > 0$ was arbitrary, we have $(c \cdot d_{\Gamma, h})(\rho') \geq (c \cdot d_{\Gamma, h})(\rho)$. \square

The second consequence is a decomposition property of **Multifac** $[\Gamma, h; V, c]$. For an orbit Q , define $(c \cdot d_{\Gamma/Q, h}) : (V_{\Gamma/Q})^n \rightarrow \mathbf{R}$ by

$$(c \cdot d_{\Gamma/Q, h})(\rho) := \sum_{s \in V_\Gamma} \sum_{1 \leq j \leq n} c(sj) d_{\Gamma/Q, h}(s/Q, \rho_j) + \sum_{1 \leq i < j \leq n} c(ij) d_{\Gamma/Q, h}(\rho_i, \rho_j) \\ (\rho = (\rho_1, \rho_2, \dots, \rho_n) \in (V_{\Gamma/Q})^n).$$

See Section 2.2.1 for the definition of Γ/Q . Consider the following problem on Γ/Q :

$$(5.1) \quad \text{Minimize } (c \cdot d_{\Gamma/Q,h})(\rho) \text{ over all locations } \rho \in (V_{\Gamma/Q})^n.$$

The optimal value of (5.1) is denoted by $\tau_Q(\Gamma, h; V, c)$, whereas the optimal value of the original problem **Multifac** $[\Gamma, h; V, c]$ is denoted by $\tau(\Gamma, h; V, c)$. Then we have

$$(5.2) \quad \tau(\Gamma, h; V, c) \geq \sum_{Q:\text{orbit}} h_Q \tau_Q(\Gamma, 1; V, c).$$

Indeed, for any optimal location ρ in **Multifac** $[\Gamma, h; V, c]$, define a location ρ/Q for Γ/Q by $\rho/Q := (\rho_1/Q, \rho_2/Q, \dots, \rho_n/Q)$. By (2.4) we have

$$(5.3) \quad \tau(\Gamma, h; V, c) = (c \cdot d_{\Gamma,h})(\rho) = \sum_{Q:\text{orbit}} h_Q (c \cdot d_{\Gamma,1})(\rho/Q) \geq \sum_{Q:\text{orbit}} h_Q \tau_Q(\Gamma, 1; V, c).$$

Note that problems **Multifac** $[\Gamma, h; V, c]$ and (5.1) can be considered for a possibly nonorientable modular graph, and the inequality relation (5.2) still holds; see [34]. A modular graph Γ is said to be *orbit-additive* if (5.2) holds in equality. Karzanov [34, Section 6] conjectured that every orientable modular graph is orbit-additive. We can solve this conjecture affirmatively.

Theorem 5.4. *Every orientable modular graph is orbit-additive.*

Proof. Take an optimal solution ρ in **Multifac** $[\Gamma, 1; V, c]$. By Theorem 5.3, ρ is also optimal to **Multifac** $[\Gamma, 1_Q; V, c]$ for every orbit Q , where 1_Q is the orbit-invariant function taking 1 on Q and 0 on $E_\Gamma \setminus Q$. Here **Multifac** $[\Gamma, 1_Q; V, c]$ is equivalent to (5.1). Hence the inequality in (5.3) holds in equality. \square

Remark 5.5. If problem (5.1) is solvable in (strongly) polynomial time for each orbit, then by Theorem 5.4 we can evaluate τ in (strongly) polynomial time, and hence **0-Ext** $[\Gamma]$ is solvable in (strongly) polynomial time by the variable-fixing technique (see the augment after Theorem 2.1). As was suggested by Karzanov [34], this approach is applicable to the case where each orbit graph of Γ is a frame. Then (5.1) is a 0-extension problem on a frame, is solvable in strongly polynomial time, and hence **0-Ext** $[\Gamma]$ is solvable in strongly polynomial time; the (strong polytime) tractability of this class of orientable modular graphs was conjectured by [34]. It should be noted that our proof of the main theorem gives only a weakly polynomial time algorithm.

Proof of Theorem 5.2. (1) \Leftrightarrow (3) and (2) \Leftrightarrow (4) follow from Theorems 4.11 and 5.1. (4) \Rightarrow (5) is obvious.

(3) \Rightarrow (5). Suppose that ρ' is a Q -neighbor of ρ . Then $\rho'/R = \rho/R$ for orbit R different from Q , and $d_{\Gamma/Q,h} = h_Q d_{\Gamma/Q,1}$. By (2.4) we have

$$\begin{aligned} 0 &\leq (c \cdot d_{\Gamma,h})(\rho') - (c \cdot d_{\Gamma,h})(\rho) = \sum_{R:\text{orbit}} (c \cdot d_{\Gamma/R,h})(\rho'/R) - (c \cdot d_{\Gamma/R,h})(\rho/R) \\ &= (c \cdot d_{\Gamma/Q,h})(\rho'/Q) - (c \cdot d_{\Gamma/Q,h})(\rho/Q) \\ &= h_Q \{ (c \cdot d_{\Gamma/Q,1})(\rho'/Q) - (c \cdot d_{\Gamma/Q,1})(\rho/Q) \} \\ &= h_Q \sum_{R:\text{orbit}} (c \cdot d_{\Gamma/R,1})(\rho'/R) - (c \cdot d_{\Gamma/R,1})(\rho/R) \\ &= h_Q \{ (c \cdot d_{\Gamma,1})(\rho') - (c \cdot d_{\Gamma,1})(\rho) \}. \end{aligned}$$

(5) \Rightarrow (3) and (5) \Rightarrow (4). By Lemma 2.16 (3), $\mathcal{L}_p^+|Q$ is convex in \mathcal{L}_p^+ , and we can define $q|Q \in \mathcal{L}_p^+|Q$ for $q \in \mathcal{L}_p^+$.

Lemma 5.6. For $p, q \in V_\Gamma$, $p' \in \mathcal{L}_p^+$, and $q' \in \mathcal{L}_q^+$, we have

$$d_{\Gamma,h}(p', q') - d_{\Gamma,h}(p, q) = \sum_{Q:\text{orbit}} h_Q \{d_{\Gamma,1}(p'|Q, q'|Q) - d_{\Gamma,1}(p, q)\}.$$

Proof. First we remark

$$(p'|Q)/R = \begin{cases} p'/Q & \text{if } Q = R, \\ p/R & \text{otherwise,} \end{cases} \quad (p' \in \mathcal{L}_p^+, Q, R : \text{orbits}).$$

Indeed, if $Q \neq R$, then $p'|Q$ and p are joined by edges in $Q \subseteq E_\Gamma \setminus R$, and hence $(p'|Q)/R = p/R$. If $Q = R$, then p' and $p'|Q$ are joined by $E_\Gamma \setminus Q$ (Lemma 2.16 (3)), and hence $p'/Q = (p'|Q)/Q$. Thus we have

$$\begin{aligned} & \sum_{Q:\text{orbit}} h_Q \{d_{\Gamma,1}(p'|Q, q'|Q) - d_{\Gamma,1}(p, q)\} \\ &= \sum_{Q:\text{orbit}} h_Q \sum_{R:\text{orbit}} \{d_{\Gamma/R,1}((p'|Q)/R, (q'|Q)/R) - d_{\Gamma/R,1}(p/R, q/R)\} \\ &= \sum_{Q:\text{orbit}} h_Q \{d_{\Gamma/Q,1}(p'/Q, q'/Q) - d_{\Gamma/Q,1}(p/Q, q/Q)\} \\ &= d_{\Gamma,h}(p', q') - d_{\Gamma,h}(p, q). \end{aligned}$$

□

Let ρ be a location, and let ρ' be a forward neighbor of ρ . For each orbit Q , define the forward Q -neighbor $\rho'|Q$ by $(\rho'|Q)_i := \rho'_i|Q$ for $i = 1, 2, \dots, n$. By the above lemma, we have

$$(c \cdot d_{\Gamma,h})(\rho') - (c \cdot d_{\Gamma,h})(\rho) = \sum_{Q:\text{orbit}} h_Q \{(c \cdot d_{\Gamma,1})(\rho'|Q) - (c \cdot d_{\Gamma,1})(\rho)\}.$$

From this, we have (5) \Rightarrow (3) and (5) \Rightarrow (4).

5.3 Proof of the main theorem (Theorem 1.6)

In this section, we complete the proof of the main theorem (Theorem 1.6) stating that **0-Ext** $[\Gamma]$ for every orientable modular graph Γ can be solved in polynomial time. By Theorem 5.1, an instance **Multifac** $[\Gamma, 1; V, c]$ of **0-Ext** $[\Gamma]$ is the problem of minimizing the sum $(c \cdot d_{\Gamma,1})$ of L-convex constraints on $\mathbf{\Gamma}^n$, where the arity of each constraint is 2. Hence, every location $\rho = (\rho_1, \rho_2, \dots, \rho_n)$, and sign $s \in \{-, +\}$, $(c \cdot d_{\Gamma,1})$ is the sum of arity-2 submodular constraints on $\mathcal{L}_\rho^s = \mathcal{L}_{\rho_1}^s \times \mathcal{L}_{\rho_2}^s \times \dots \times \mathcal{L}_{\rho_n}^s$, where each semilattice $\mathcal{L}_{\rho_i}^s = \mathcal{L}_{\rho_i}^s(\mathbf{\Gamma})$ is constructed in polynomial time (Proposition 4.1). By Theorem 3.9, we can minimize $(c \cdot d_{\Gamma,1})$ over $\mathcal{L}_\rho^+ \cup \mathcal{L}_\rho^-$ in polynomial time. Therefore we can assume that we have a *descent oracle*, an oracle that returns an optimal solution of this (local) problem.

By Theorem 4.11 and the steepest descent algorithm, we can obtain a global optimal solution. As mentioned already, we do not know whether the number of descent steps is polynomially bounded. Fortunately, in the case of multifacility location functions, a cost-scaling approach gives a *weakly* polynomial bound on the number of descent steps. Now the main theorem (Theorem 1.6) follows from the following.

Proposition 5.7. Suppose that c is integer-valued. **Multifac** $[\Gamma, 1; V, c]$ can be solved with $O(|V|^2 \text{diam } \Gamma \log C)$ calls of the descent oracle, where $C := \max\{c(xy) \mid xy \in \binom{V}{2}\}$ and $\text{diam } \Gamma$ denotes the diameter of Γ .

Proof. Let $\lfloor c/2 \rfloor : \binom{V}{2} \rightarrow \mathbf{Z}_+$ be defined by $\lfloor c/2 \rfloor(xy) := \lfloor c(xy)/2 \rfloor$ (the largest integer not exceeding $c(xy)/2$). Let ρ be an optimal location in $\mathbf{Multifac}[G, 1; V, \lfloor c/2 \rfloor]$, and let ρ^* be an optimal location in $\mathbf{Multifac}[G, 1; V, c]$. We show

$$(5.4) \quad (c \cdot d_{G,1})(\rho) - (c \cdot d_{G,1})(\rho^*) \leq |V|^2 \text{diam } G.$$

If (5.4) is true, then the number of the descent steps from an initial starting point ρ is bounded by $|V|^2 \text{diam } G$. Consequently, by recursive scaling, we obtain an optimal solution for $\mathbf{Multifac}[G, 1; V, c]$ in $O(|V|^2 \text{diam } G \log C)$ descent steps.

Let ϵ be a $\{0, 1\}$ -valued cost defined by $\epsilon := c - 2\lfloor c/2 \rfloor$. Then we have

$$\begin{aligned} & (c \cdot d_{G,1})(\rho) - (c \cdot d_{G,1})(\rho^*) \\ &= (2\lfloor c/2 \rfloor \cdot d_{G,1})(\rho) - (2\lfloor c/2 \rfloor \cdot d_{G,1})(\rho^*) + (\epsilon \cdot d_{G,1})(\rho) - (\epsilon \cdot d_{G,1})(\rho^*) \\ &\leq (n|V_G| + n(n-1)/2) \text{diam } G, \end{aligned}$$

where we use the facts that ρ is also an optimal location in $\mathbf{Multifac}[G, 1; V, 2\lfloor c/2 \rfloor]$ and that each term in $(\epsilon \cdot d_{G,1})(\rho')$ (for any location ρ') is at most $\text{diam } G$. \square

5.4 Minimum 0-extension problems for metrics

Let μ be a metric on a finite set S (not necessarily a graph metric). We can naturally consider the minimum 0-extension problem $\mathbf{0-Ext}[\mu]$ for a general μ formulated as: *Given a set $V \supseteq S$ and $c : \binom{V}{2} \rightarrow \mathbf{Q}_+$, find a 0-extension (V, d) of (S, μ) with $\sum_{xy} c(xy)d(x, y)$ minimum.*

Metric μ is said to be *modular* if (S, μ) is a modular metric space (see Section 2). Let H_μ be the graph on the vertex set S with edge set E_{H_μ} given as: $xy \in E_{H_\mu} \Leftrightarrow$ there is no $z \in S \setminus \{x, y\}$ with $\mu(x, z) + \mu(z, y) = \mu(x, y)$. H_μ is called the *support graph* of μ . Karzanov [35] extended the hardness result (Theorem 1.5) to the following.

Theorem 5.8 ([35]). *If μ is not modular or H_μ is not orientable, then $\mathbf{0-Ext}[\mu]$ is NP-hard.*

We can also consider LP-relaxation $\mathbf{Ext}[\mu]$ obtained by relaxing 0-extensions into extensions in $\mathbf{0-Ext}[\mu]$. Extending Theorem 1.3, Bandelt, Chepoi, and Karzanov [4] proved that $\mathbf{Ext}[\mu]$ is exact if and only if μ is modular and H_μ is frame.

Our framework covers $\mathbf{0-Ext}[\mu]$ for a metric μ such that μ is modular and H_μ is orientable. Indeed μ induces the edge-length $\bar{\mu}$ on H_μ by $\bar{\mu}(pq) = \mu(p, q)$ ($pq \in E_{H_\mu}$). From the definition of the support graph H_μ , we have $\mu = d_{H_\mu, \bar{\mu}}$. Moreover, it was shown in [2] (see [34, Section 2]) that

$$(5.5) \quad \text{if } \mu \text{ is modular, then } H_\mu \text{ is a modular graph, and } \bar{\mu} \text{ is orbit-invariant.}$$

Hence we can apply the argument in Section 5 to $\mathbf{Multifac}[H_\mu, \bar{\mu}; V, c]$ to obtain results for $\mathbf{0-Ext}[\mu]$. By Theorems 1.6 and 5.2, we obtain the converse of Theorem 5.8, which completes the classification of those metrics for which $\mathbf{0-Ext}[\mu]$ is tractable.

Theorem 5.9. *If μ is modular and H_μ is orientable, then $\mathbf{0-Ext}[\mu]$ is solvable in polynomial time.*

6 Concluding remark

In this paper, we established the ‘‘P or NP-hard’’ classification of the minimum 0-extension problem. This dichotomy result is related to a special case of a dichotomy

theorem for *finite-valued CSP* due to Thapper and Živný [57]. Here we briefly explain their result and its relation to our result.

To describe their result, we formulate valued CSP in a setting slightly different from that in Section 2.1. Let D be a finite set. A (finite-valued) *cost function* on D is a function $f : D^k \rightarrow \mathbf{Q}$ for some $k = k_f$. A (finite-valued) *constraint language*, or simply, *language* on D is a set Λ of cost functions on D . Let x_1, x_2, \dots, x_n be a set of variables. By a Λ -*constraint*, we here mean a triple (w, f, σ) of a nonnegative weight $w \in \mathbf{Q}_+$, a cost function f in Λ , and a map $\sigma : \{1, 2, \dots, k_f\} \rightarrow \{1, 2, \dots, n\}$. For a constraint language Λ , the problem $\mathbf{VCSP}[\Lambda]$ is formulated as:

VCSP $[\Lambda]$: Given a set \mathcal{C} of Λ -constraints,

$$\text{minimize } \sum_{(w,f,\sigma) \in \mathcal{C}} w f(x_{\sigma(1)}, x_{\sigma(2)}, \dots, x_{\sigma(k_f)}) \text{ over all } x = (x_1, x_2, \dots, x_n) \in D^n.$$

This is also a subclass of **VCSP** studied in Section 2.1. Therefore we can consider the basic LP relaxation (BLP). An m -ary *fractional polymorphism* for Λ is a formal convex combination $\omega = \sum_{\vartheta} \omega(\vartheta) \vartheta$ of m -ary operations $\vartheta : D^m \rightarrow D$ such that

$$\begin{aligned} \frac{1}{m} \{f(x^1) + f(x^2) + \dots + f(x^m)\} &\geq \sum_{\vartheta} \omega(\vartheta) f(\vartheta(x^1, x^2, \dots, x^m)) \\ &\quad (f \in \Lambda, x^1, x^2, \dots, x^m \in D^{k_f}), \end{aligned}$$

where an operation $\vartheta : D \times D \times \dots \times D \rightarrow D$ is extended to an operation $D^k \times D^k \times \dots \times D^k \rightarrow D^k$ by $(\vartheta(x^1, x^2, \dots, x^k))_i := \vartheta(x_i^1, x_i^2, \dots, x_i^k)$. If $m = 1$, then ω is said to be *unary*, and if $m = 2$, then ω is said to be *binary*. A binary fractional polymorphism ω is said to be *symmetric* if its support consists of symmetric operations, i.e., operations ϑ satisfy $\vartheta(x, y) = \vartheta(y, x)$ for $x, y \in D$, and ω is said to be *idempotent* if its support consists of idempotent operations, i.e., operations ϑ satisfy $\vartheta(x, x) = x$ for $x \in D$.

Just after the developments [39, 56] (see [41]), Thapper and Živný [57] established the following dichotomy theorem for finite-valued CSP. Here a language Λ is said to be a *core* if for every unary fractional polymorphism, its support consists of injective operations; It is shown in [57] that $\mathbf{VCSP}[\Lambda]$ is polynomial time reducible to $\mathbf{VCSP}[\Lambda']$ for a core language Λ' .

Theorem 6.1 ([57]). *Let Λ be a finite-valued core language on D . If Λ admits a binary symmetric and idempotent fractional polymorphism, then BLP is exact for Λ , and $\mathbf{VCSP}[\Lambda]$ can be solved in polynomial time. Otherwise $\mathbf{VCSP}[\Lambda]$ is NP-hard.*

The minimum 0-extension problem can naturally be formulated as $\mathbf{VCSP}[\Lambda]$. Let μ be a (rational-valued) metric on D ; we can assume that $\mu(s, t) > 0$ whenever $s \neq t$. The metric $\mu : D^2 \rightarrow \mathbf{Q}_+$ itself is regarded as a binary cost function on D . For $s \in D$, let $\mu_s : D \rightarrow \mathbf{Q}_+$ be a unary cost function on D defined by

$$\mu_s(x) := \mu(s, x) \quad (x \in D).$$

Let Λ_μ be a (finite-valued) constraint language defined by

$$\Lambda_\mu := \{\mu\} \cup \{\mu_s\}_{s \in D}.$$

By definition, we have:

Lemma 6.2. $\mathbf{VCSP}[\Lambda_\mu] = \mathbf{0-Ext}[\mu]$.

Moreover we have:

Lemma 6.3. Λ_μ is a core language.

Proof. Take a unary fractional polymorphism $\omega = \sum_{\vartheta} \omega(\vartheta)\vartheta$ for Λ_μ . Then it holds

$$0 = \mu(s, s) \geq \sum_{\vartheta} \omega(\vartheta)\mu(s, \vartheta(s)) \geq 0.$$

Therefore $\vartheta(s) = s$ must hold. This means that the support of ω consists of the identity map on D , which is trivially injective. \square

Therefore Theorem 6.1 is applicable to **0-Ext** $[\mu]$. In particular, Theorems 5.8 and 5.9 can be viewed as a sharpening of Theorem 6.1 for constraint languages Λ_μ . For those metrics μ in Theorem 5.9, core language Λ_μ must have a binary symmetric and idempotent fractional polymorphism, and **0-Ext** $[\mu]$ must be solved directly by BLP (under the assumption $P \neq NP$). In Remark 3.10, we have verified this fact for the case where μ is the metric on a modular semilattice \mathcal{L} ; it is a good exercise to construct a binary symmetric and idempotent fractional polymorphism from (3.11). We however could not find such a fractional polymorphism for the general case. As seen in Section 3, a fractional polymorphism can rather be complicated and consist of a large number of operations. To construct a fractional polymorphism as required, it might need a further thorough investigation on orientable modular graphs.

Acknowledgments

We thank the referee for helpful comments, and thank Kazuo Murota for careful reading and numerous helpful comments, Kei Kimura for discussion on Valued-CSP, Satoru Iwata for communicating the paper [43] of Kuivinen, Akiyoshi Shioura for the paper [38] of Kolmogorov, and Satoru Fujishige for the paper [27] of Huber-Kolmogorov. This research is partially supported by the Aihara Project, the FIRST program from JSPS, by Global COE Program “The research and training center for new development in mathematics” from MEXT, and by a Grant-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

References

- [1] H.-J. Bandelt, Networks with Condorcet solutions, *European Journal of Operational Research* **20** (1985), 314–326.
- [2] H.-J. Bandelt, Hereditary modular graphs, *Combinatorica* **8** (1988), 149–157.
- [3] H.-J. Bandelt and V. Chepoi, Metric graph theory and geometry: a survey, in: J.E. Goodman, J. Pach, and R. Pollack eds., *Surveys on discrete and computational geometry: Twenty Years Later*, 49–86, American Mathematical Society, Providence, 2008.
- [4] H.-J. Bandelt, V. Chepoi, and A. V. Karzanov, A characterization of minimizable metrics in the multifacility location problem, *European Journal of Combinatorics* **21** (2000), 715–725.
- [5] H.-J. Bandelt, M. van de Vel, and E. Verheul, Modular interval spaces, *Mathematische Nachrichten* **163** (1993) 177–201.
- [6] G. Birkhoff, *Lattice Theory*, American Mathematical Society, New York, 1940; 3rd edn., American Mathematical Society, Providence, RI, 1967.
- [7] S. Bistarelli, U. Montanari, and F. Rossi, Semiring-based constraint satisfaction and optimization, *Journal of the ACM* **44** (1997), 201–236.

- [8] A. Bouchet, Multimatroids. I. Coverings by independent sets, *SIAM Journal on Discrete Mathematics* **10** (1997), 626–646.
- [9] R. Chandrasekaran and S. N. Kabadi, Pseudomatroids, *Discrete Mathematics* **71** (1988), 205–217.
- [10] J. Chalopin, V. Chepoi, H. Hirai, and D. Osajda, Weakly modular graphs and nonnegative curvature, preprint, 2014, [arXiv:1409.3892](#).
- [11] V. Chepoi, Classification of graphs by means of metric triangles, *Metody Diskretnogo Analiza* **49** (1989), 75–93 (in Russian).
- [12] V. Chepoi, A multifacility location problem on median spaces, *Discrete Applied Mathematics* **64** (1996) 1–29.
- [13] V. Chepoi, Graphs of some CAT(0) complexes, *Advances in Applied Mathematics* **24** (2000), 125–179.
- [14] E. Dahlhaus, D. S. Johnson, C. H. Papadimitriou, P. D. Seymour, and M. Yannakakis, The complexity of multiterminal cuts, *SIAM Journal on Computing* **23** (1994), 864–894.
- [15] A. W. M. Dress and R. Scharlau, Gated sets in metric spaces, *Aequationes Mathematicae* **34** (1987), 112–120.
- [16] S. Fujishige, *Submodular Functions and Optimization, 2nd Edition*, Elsevier, Amsterdam, 2005.
- [17] S. Fujishige and K. Murota, Notes on L-/M-convex functions and the separation theorems, *Mathematical Programming, Series A* **88** (2000), 129–146.
- [18] S. Fujishige and S. Tanigawa, A min-max theorem for transversal submodular functions and its implications, *SIAM Journal on Discrete Mathematics*, to appear.
- [19] S. Fujishige, S. Tanigawa, and Y. Yoshida, Generalized skew bisubmodularity: A characterization and a min-max theorem, *Discrete Optimization* **12** (2014), 1–9.
- [20] M. Grötschel, L. Lovász, and A. Schrijver, *Geometric Algorithms and Combinatorial Optimization*, Springer-Verlag, Berlin, 1988.
- [21] H. Hirai, Tight spans of distances and the dual fractionality of undirected multiflow problems, *Journal of Combinatorial Theory, Series B* **99** (2009), 843–868.
- [22] H. Hirai, Folder complexes and multiflow combinatorial dualities, *SIAM Journal on Discrete Mathematics* **25** (2011), 1119–1143.
- [23] H. Hirai, Half-integrality of node-capacitated multiflows and tree-shaped facility locations on trees, *Mathematical Programming, Series A* **137** (2013), 503–530.
- [24] H. Hirai, The maximum multiflow problems with bounded fractionality, *Mathematics of Operations Research* **39** (2014), 60–104.
- [25] H. Hirai, Discrete convexity for multiflows and 0-extensions, in: *Proceeding of 8th Japanese-Hungarian Symposium on Discrete Mathematics and Its Applications*, 2013, pp. 209–223.
- [26] H. Hirai, L-convexity on graph structures, in preparation.
- [27] A. Huber and V. Kolmogorov, Towards minimizing k -submodular functions, in: *Proceedings of the 2nd International Symposium on Combinatorial Optimization (ISCO'12)*, LNCS 7422, Springer, Berlin, 2012, pp. 451–462.
- [28] A. Huber and A. Krokhin, Oracle tractability of skew bisubmodular functions, *SIAM Journal on Discrete Mathematics* **28** (2014), 1828–1837.

- [29] A. Huber, A. Krokhin, and R. Powell, Skew bisubmodularity and valued CSPs, *SIAM Journal on Computing* **43** (2014), 1064–1084.
- [30] S. Iwata, L. Fleischer, and S. Fujishige, A combinatorial strongly polynomial algorithm for minimizing submodular functions, *Journal of the ACM* **48** (2001), 761–777.
- [31] A. V. Karzanov, Polyhedra related to undirected multicommodity flows, *Linear Algebra and its Applications* **114/115** (1989), 293–328.
- [32] A. V. Karzanov, Minimum 0-extensions of graph metrics, *European Journal of Combinatorics* **19** (1998), 71–101.
- [33] A. V. Karzanov, Metrics with finite sets of primitive extensions, *Annals of Combinatorics* **2** (1998), 211–241.
- [34] A. V. Karzanov, One more well-solved case of the multifacility location problem, *Discrete Optimization* **1** (2004), 51–66.
- [35] A. V. Karzanov, Hard cases of the multifacility location problem, *Discrete Applied Mathematics* **143** (2004), 368–373.
- [36] J. Kleinberg and É. Tardos, Approximation algorithms for classification problems with pairwise relationships: metric labeling and Markov random fields, *Journal of the ACM* **49** (2002), 616–639.
- [37] A. W. J. Kolen, *Tree Network and Planar Rectilinear Location Theory*, CWI Tract 25, Center for Mathematics and Computer Science, Amsterdam, 1986.
- [38] V. Kolmogorov, Submodularity on a tree: Unifying L^1 -convex and bisubmodular functions, in: *Proceedings of the 36th International Symposium on Mathematical Foundations of Computer Science (MFCS'11)*, LNCS 6907, Springer, Berlin, 2011, pp. 400–411.
- [39] V. Kolmogorov, The power of linear programming for finite-valued CSPs: A constructive characterization, in: *Proceedings of the 40th International Colloquium, ICALP 2013*, LNCS 7965, 2013, pp. 625–636.
- [40] V. Kolmogorov and A. Shioura, New algorithms for convex cost tension problem with application to computer vision, *Discrete Optimization* **6** (2009), 378–393.
- [41] V. Kolmogorov, J. Thapper, and S. Živný, The power of linear programming for general-valued CSPs, preprint, 2013, [arXiv:1311.4219](https://arxiv.org/abs/1311.4219).
- [42] V. Kolmogorov and S. Živný, The complexity of conservative valued CSPs, *Journal of the ACM* **60** (2013), Article No. 10.
- [43] F. Kuivinen, On the complexity of submodular function minimisation on diamonds, *Discrete Optimization* **8** (2011), 459–477.
- [44] L. Lovász, Submodular functions and convexity, in: A. Bachem, M. Grötschel, and B. Korte, eds., *Mathematical Programming—The State of the Art*, Springer-Verlag, Berlin, 1983, 235–257.
- [45] K. Murota, Discrete convex analysis, *Mathematical Programming* **83** (1998), 313–371.
- [46] K. Murota, Algorithms in discrete convex analysis, *IEICE Transactions on Systems and Information*, **E83-D** (2000), 344–352.
- [47] K. Murota, *Discrete Convex Analysis*, SIAM, Philadelphia, 2003.
- [48] K. Murota and A. Shioura, M-convex function on generalized polymatroid, *Mathematics of Operations Research* **24** (1999), 95–105.

- [49] K. Murota and A. Tamura, Proximity theorems of discrete convex functions, *Mathematical Programming, Series A* **99** (2004), 539–562.
- [50] M. Nakamura, A characterization of greedy sets: Universal polymatroids (I), *Scientific Papers of College of Arts and Science, The University of Tokyo* **38** (1988), 155–167.
- [51] J. C. Picard and D. H. Ratliff, A cut approach to the rectilinear distance facility location problem, *Operations Research* **26** (1978), 422–433.
- [52] L. Qi, Directed submodularity, ditroids and directed submodular flows, *Mathematical Programming* **42** (1988), 579–599.
- [53] T. Schiex, H. Fargier, and G. Verfaillie, Valued constraint satisfaction problems: hard and easy problems, In *Proceedings of the 14th International Joint Conference on Artificial Intelligence (IJCAI'95)*, 1995.
- [54] A. Schrijver, A combinatorial algorithm minimizing submodular functions in strongly polynomial time, *Journal of Combinatorial Theory, Series B* **80** (2000), 346–355.
- [55] B. C. Tansel, R. L. Francis, and T. J. Lowe, Location on networks I, II, *Management Science* **29** (1983), 498–511.
- [56] J. Thapper and S. Živný, The power of linear programming for valued CSPs, in: *Proceedings of the 53rd Annual IEEE Symposium on Foundations of Computer Science (FOCS'12)*, 2012, pp. 669–678.
- [57] J. Thapper and S. Živný, The complexity of finite-valued CSPs, in: *Proceedings of the 45th ACM Symposium on the Theory of Computing (STOC'13)*, 2013, pp. 695–704.
- [58] M. L. J. van de Vel, *Theory of Convex Structures*, North-Holland, Amsterdam, 1993.
- [59] S. Živný, *The Complexity of Valued Constraint Satisfaction Problems*, Springer, Heidelberg, 2012.